GenCore version 5.1.9
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 3, 2006, 20:51:16 ; Search time 2057.56 Seconds
(without alignments)
12719.092 Million cell updates/sec

Title: US-10-764-316-5
Perfect score: 468
Sequence: 1 tgaaaataaatggaacacac.....tggatgattgcaggagattt 468

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 48236798 seqs, 27959665780 residues

Word size : 1

Total number of hits satisfying chosen parameters: 96473154

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est3:*
3: gb_est4:*
4: gb_est5:*
5: gb_est6:*
6: gb_htc:*
7: gb_est2:*
8: gb_est7:*
9: gb_est8:*
10: gb_est9:*
11: gb_gss1:*
12: gb_gss2:*
13: gb_gss3:*
14: gb_gss4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	21	4.5	627	13	CW375149	CW375149 fsbb001f0
c	2	21	4.5	649	13	CW258112	CW258112 104_723_1
	3	21	4.5	682	13	CW375148	CW375148 fsbb001f0
c	4	21	4.5	983	14	CT013281	CT013281 KBrH121I1
	5	21	4.5	999	14	DU750205	DU750205 ASNF1346.
c	6	21	4.5	1217	12	CC314499	CC314499 TAM32-17E
	7	20	4.3	245	13	CL219260	CL219260 ZMMBBc042
c	8	20	4.3	611	14	DX327753	DX327753 OR_ABa024
	9	20	4.3	613	9	CX687527	CX687527 ydd80c08.
	10	20	4.3	626	9	CX688942	CX688942 yde22f05.
c	11	20	4.3	640	10	DT673917	DT673917 s13dLT65C
c	12	20	4.3	679	11	BH557567	BH557567 BOGOG92TF
c	13	20	4.3	785	11	BZ162280	BZ162280 CH230-276
	14	20	4.3	896	3	BU795670	BU795670 SJF2DUA01
c	15	19	4.1	230	7	BB212456	BB212456 BB212456
c	16	19	4.1	281	7	BB369985	BB369985 BB369985
c	17	19	4.1	288	3	BQ803876	BQ803876 WHE2843_B
c	18	19	4.1	415	2	BG264482	BG264482 daa82f12.
	19	19	4.1	422	12	CC812860	CC812860 ZMMBBc052
	20	19	4.1	427	11	BH752910	BH752910 SALK_0194
c	21	19	4.1	428	11	AQ514846	AQ514846 HS_5199_A
	22	19	4.1	437	8	CN184594	CN184594 UCRCS04_0
	23	19	4.1	451	11	BZ379446	BZ379446 SALK_1133
	24	19	4.1	458	7	AW680443	AW680443 WS1_5_B12
c	25	19	4.1	468	9	CX538294	CX538294 s13dNF41E
	26	19	4.1	478	11	AQ130921	AQ130921 HS_3027_B
	27	19	4.1	483	10	DV773477	DV773477 McClintoc
c	28	19	4.1	492	14	CR900936	CR900936 Sus scrof
c	29	19	4.1	522	7	BE429624	BE429624 TAS002.B0
c	30	19	4.1	549	2	BM136116	BM136116 WHE2602_H
c	31	19	4.1	558	13	CZ891668	CZ891668 fgha001f0
c	32	19	4.1	561	13	CW484121	CW484121 fsbb001f2
c	33	19	4.1	567	4	CA639214	CA639214 wre1n.pk0
c	34	19	4.1	577	4	CB085812	CB085812 hg17a01.g
c	35	19	4.1	577	14	DX171753	DX171753 OR_ABa003
	36	19	4.1	577	14	CR488931	CR488931 mth2-159E
	37	19	4.1	578	3	BQ400766	BQ400766 NISC_mp11
c	38	19	4.1	579	11	AZ704872	AZ704872 RPCI-23-2
c	39	19	4.1	587	12	CE742874	CE742874 tigr-gss-
c	40	19	4.1	595	11	AZ704151	AZ704151 RPCI-23-2
c	41	19	4.1	600	14	DX172123	DX172123 OR_ABa003
c	42	19	4.1	617	13	DU224130	DU224130 109857415
c	43	19	4.1	619	7	AW641002	AW641002 cm02c07.w
c	44	19	4.1	624	7	BE492287	BE492287 WHE0557_B
c	45	19	4.1	624	12	CE178051	CE178051 tigr-gss-

ALIGNMENTS

RESULT 1
CW375149/c
LOCUS CW375149 627 bp DNA linear GSS 01-NOV-2004
DEFINITION fsbb001f052g07k0 Sorghum methylation filtered library (LibID: 104)
Sorghum bicolor genomic clone fsbb001f052g07, genomic survey sequence.
ACCESSION CW375149
VERSION CW375149.1 GI:55093593
KEYWORDS GSS.
SOURCE Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE 1 (bases 1 to 627)
AUTHORS Bedell,J.A., Budiman,M.A., Nunberg,A., Citek,R.W., Robbins,D., Jones,J., Flick,E., Rohlfing,T., Fries,J., Bradford,K., McMenamy,J., Smith,M., Holeman,H., Roe,B.A, Wiley,G., Korf,I.F., Rabinowicz,P.D., Lakey,N., McCombie,W.R., Jeddeloh,J.A. and Martienssen,R.A.
TITLE Sorghum genome sequencing by methylation filtration
JOURNAL PLoS Biol. 3 (1), e13 (2005)
PUBMED 15660154
COMMENT Contact: Bedell JA
Orion Genomics, LLC
4041 Forest Park Ave, St. Louis, MO 63108, USA
Tel: 314 615 6979
Fax: 314 615 5975
Email: jbedell@oriongenomics.com
Plate: fsbb001f052 row: g column: 07
Seq primer: k Reverse
Class: methylation filtered
High quality sequence stop: 627.
FEATURES Location/Qualifiers
source 1. .627
/organism="Sorghum bicolor"
/mol_type="genomic DNA"
/cultivar="ATx623"
/db_xref="taxon:4558"
/clone="fsbb001f052g07"
/clone_lib="Sorghum methylation filtered library (LibID: 104)"

/note="Organ: leaf; Vector: pBCSK(-); Site_1: HincII; DNA prepared from purified nuclei was randomly sheared, end-repaired, size fractionated to enrich for the 0.5 to 5 kb fraction, ligated into HincII-digested pBCSK(-) vector and electroporated into E. coli cells. This is a methylation filtered library."

ORIGIN

Query Match 4.5%; Score 21; DB 13; Length 627;
Best Local Similarity 100.0%; Pred. No. 7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGAAAATAAATGGAACACACA 21
||||||||||||||||||||
Db 467 TGAAAATAAATGGAACACACA 447

RESULT 2
CW258112/c
LOCUS CW258112 649 bp DNA linear GSS 30-OCT-2004
DEFINITION 104_723_11227699_116_35158_066 Sorghum methylation filtered library (LibID: 104) Sorghum bicolor genomic clone 11227699, genomic survey sequence.
ACCESSION CW258112
VERSION CW258112.1 GI:54966939
KEYWORDS GSS.
SOURCE Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE 1 (bases 1 to 649)
AUTHORS Bedell,J.A., Budiman,M.A., Nunberg,A., Citek,R.W., Robbins,D., Jones,J., Flick,E., Rohlfing,T., Fries,J., Bradford,K., McMenamy,J., Smith,M., Holeman,H., Roe,B.A, Wiley,G., Korf,I.F., Rabinowicz,P.D., Lakey,N., McCombie,W.R., Jeddeloh,J.A. and Martienssen,R.A.
TITLE Sorghum genome sequencing by methylation filtration
JOURNAL PLoS Biol. 3 (1), e13 (2005)
PUBMED 15660154
COMMENT Contact: Bedell JA
Orion Genomics, LLC
4041 Forest Park Ave, St. Louis, MO 63108, USA
Tel: 314 615 6979
Fax: 314 615 5975
Email: jbedell@oriongenomics.com
Plate: 723 row: o column: 19
Seq primer: T3 Reverse
Class: methylation filtered
High quality sequence stop: 649.
FEATURES Location/Qualifiers
source 1. .649
/organism="Sorghum bicolor"
/mol_type="genomic DNA"
/cultivar="ATx623"
/db_xref="taxon:4558"
/clone="11227699"
/clone_lib="Sorghum methylation filtered library (LibID: 104)"
/note="Organ: leaf; Vector: pBCSK(-); Site_1: HincII; DNA prepared from purified nuclei was randomly sheared, end-repaired, size fractionated to enrich for the 0.5 to 5 kb fraction, ligated into HincII-digested pBCSK(-) vector and electroporated into E. coli cells. This is a methylation filtered library."

ORIGIN

Query Match 4.5%; Score 21; DB 13; Length 649;
Best Local Similarity 100.0%; Pred. No. 7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGAAAATAAATGGAACACACA 21
||||||||||||||||||||
Db 425 TGAAAATAAATGGAACACACA 405

RESULT 3
CW375148
LOCUS CW375148 682 bp DNA linear GSS 01-NOV-2004
DEFINITION fsbb001f052g07f0 Sorghum methylation filtered library (LibID: 104) Sorghum bicolor genomic clone fsbb001f052g07, genomic survey sequence.
ACCESSION CW375148
VERSION CW375148.1 GI:55093592
KEYWORDS GSS.
SOURCE Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE 1 (bases 1 to 682)
AUTHORS Bedell,J.A., Budiman,M.A., Nunberg,A., Citek,R.W., Robbins,D., Jones,J., Flick,E., Rohlfing,T., Fries,J., Bradford,K., McMenamy,J., Smith,M., Holeman,H., Roe,B.A, Wiley,G., Korf,I.F., Rabinowicz,P.D., Lakey,N., McCombie,W.R., Jeddeloh,J.A. and Martienssen,R.A.
TITLE Sorghum genome sequencing by methylation filtration
JOURNAL PLoS Biol. 3 (1), e13 (2005)
PUBMED 15660154
COMMENT Contact: Bedell JA
Orion Genomics, LLC
4041 Forest Park Ave, St. Louis, MO 63108, USA
Tel: 314 615 6979
Fax: 314 615 5975
Email: jbedell@oriongenomics.com
Plate: fsbb001f052 row: g column: 07
Seq primer: f Forward
Class: methylation filtered
High quality sequence stop: 682.
FEATURES Location/Qualifiers
source 1. .682
/organism="Sorghum bicolor"
/mol_type="genomic DNA"
/cultivar="ATx623"
/db_xref="taxon:4558"
/clone="fsbb001f052g07"
/clone_lib="Sorghum methylation filtered library (LibID: 104)"
/note="Organ: leaf; Vector: pBCSK(-); Site_1: HincII; DNA prepared from purified nuclei was randomly sheared, end-repaired, size fractionated to enrich for the 0.5 to 5 kb fraction, ligated into HincII-digested pBCSK(-) vector and electroporated into E. coli cells. This is a methylation filtered library."

ORIGIN

Query Match 4.5%; Score 21; DB 13; Length 682;
Best Local Similarity 100.0%; Pred. No. 7.1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGAAAATAAATGGAACACACA 21
||||||||||||||||||||
Db 477 TGAAAATAAATGGAACACACA 497

RESULT 4
CT013281/c
LOCUS CT013281 983 bp DNA linear GSS 09-AUG-2005
DEFINITION KBrH121I17 genomic clone, KBrH (HindIII) BAC library Brassica rapa subsp. pekinensis, genomic survey sequence.
ACCESSION CT013281
VERSION CT013281.1 GI:71477141
KEYWORDS GSS.
SOURCE Brassica rapa subsp. pekinensis
ORGANISM Brassica rapa subsp. pekinensis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1
  AUTHORS   Viehoever,P., Holtgraewe,D. and Weisshaar,B.
  TITLE     BAC end sequences of Brassica rapa
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 983)
  AUTHORS   Li,Y. and Weisshaar,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-AUG-2005) Weisshaar B., Bielefeld University,
            Institute for Genome Research, Universitaetsstrasse 25, D-33594
            Bielefeld, Germany
COMMENT     Contact: Bernd Weisshaar
            Bielefeld University, Institute for Genome Research
            Universitaetsstrasse 25, D-33594 Bielefeld, Germany Email:
            bernd.weisshaar@uni-bielefeld.de
            BAC end sequences of Brassica rapa BAC clone KBrH121I17; generated
            as contribution to the 'Multinational Brassica rapa Sequencing
            Project' Seq primer: sp6B ATTTAGGTGACACTATAG
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .983
                     /organism="Brassica rapa subsp. pekinensis"
                     /mol_type="genomic DNA"
                     /strain="Chiifu type 401-42"
                     /cultivar="Chiifu"
                     /sub_species="pekinensis"
                     /db_xref="taxon:51351"
                     /clone="KBrH121I17"
                     /clone_lib="KBrH, Brassica rapa HindIII BAC library
                     GF-SCF-1002, Vector: pCUGIBac1"
                     /lab_host="E.coli DH10B"
ORIGIN

  Query Match             4.5%;  Score 21;  DB 14;  Length 983;
  Best Local Similarity   100.0%;  Pred. No. 7.2;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        324 TTGGTAATGATGTTTTCTTCT 344
              |||||||||||||||||||||
Db        821 TTGGTAATGATGTTTTCTTCT 801


RESULT 5
DU750205
LOCUS       DU750205                999 bp    DNA     linear   GSS 27-JAN-2006
DEFINITION  ASNF1346.g2 HF130_10-06-02 uncultured marine microorganism
            HF130_10-06-02 genomic clone HF0130_014B01, genomic survey
            sequence.
ACCESSION   DU750205
VERSION     DU750205.1  GI:85760041
KEYWORDS    GSS.
SOURCE      uncultured marine microorganism HF130_10-06-02
  ORGANISM  uncultured marine microorganism HF130_10-06-02
            unclassified sequences; environmental samples.
REFERENCE   1  (bases 1 to 999)
  AUTHORS   DeLong,E.F., Preston,C.M., Mincer,T., Rich,V., Hallam,S.J.,
            Frigaard,N.U., Martinez,A., Sullivan,M., Edwards,R., Chisholm,S.W.
            and Karl,D.M.
  TITLE     Comparative genomics reveals ecological trends in stratified
            microbial communities in the ocean's interior
  JOURNAL   Science (2006) In press
COMMENT     Contact: Susan Lucas, Alex Copeland, Sam Pitluck, Alla Lapidus,
            Kerrie Barry, Tijana Glavinadelrio, David Bruce, Paul Richardson
            and Edward DeLong
            US DOE Joint Genome Institute
            US DOE Joint Genome Institute
            2800 Mitchell Drive B100, Walnut Creek, CA 94598-1698, USA
            Tel: 617-253-5271
            Fax: 617-253-2679
            Email: PMRichardson@lbl.gov; delong@mit.edu
            North Pacific Subtropical Gyre (Hawaii) picoplankton genomic fosmid
            DNA library prepared from marine picoplankton in the less than 1.6
            um, greater than 0.22 um fraction. Sample Date: 10/6/2002
            Coordinates: 22.45 N, 158 W Depth 130 m Temperature: 22.19 C
            Salinity: 35.31 psu Oxygen: 204.9 umol/kg
            Class: fosmid ends.
FEATURES             Location/Qualifiers
     source          1. .999
                     /organism="uncultured marine microorganism HF130_10-06-02"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:361146"
                     /clone="HF0130_014B01"
                     /cell_type="marine picoplankton, less than 1.8 um, greater
                     than 0.22 um fraction"
                     /clone_lib="HF130_10-06-02"
                     /note="Vector: pCC1FOS; North Pacific Subtropical Gyre
                     (Hawaii) picoplankton genomic fosmid DNA library prepared
                     from marine picoplankton in the less than 1.6 um, greater
                     than 0.22 um fraction.  Picoplankton collected  at 130 m
                     depth on 10/6/2002, Coordinates: 22.45 N,158 W. Sample
                     Date: 10/6/2002 Coordinates: 22.45 N, 158 W Depth 130 m
                     Temperature: 22.19 C Salinity: 35.31 psu Oxygen: 204.9
                     umol/kg"
ORIGIN

  Query Match             4.5%;  Score 21;  DB 14;  Length 999;
  Best Local Similarity   100.0%;  Pred. No. 7.2;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        254 AGAAGAGCTGATGAGGTGATT 274
              |||||||||||||||||||||
Db        261 AGAAGAGCTGATGAGGTGATT 281


RESULT 6
CC314499/c
LOCUS       CC314499                1217 bp    DNA     linear   GSS 14-MAY-2003
DEFINITION  TAM32-17E11_EC1.1 TAM32 Gallus gallus genomic clone TAM32-17E11,
            genomic survey sequence.
ACCESSION   CC314499
VERSION     CC314499.1  GI:30708554
KEYWORDS    GSS.
SOURCE      Gallus gallus (chicken)
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 1217)
  AUTHORS   Kremitzki,C., Higginbotham,J., Wylie,K., Carter,J., McPherson,J.,
            Warren,W., Graves,T., Mardis,E. and Wilson,R.
  TITLE     Gallus gallus BAC End Reads
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Richard K. Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 152000   Std Error: 0.00
            Seq primer: EC1 TACGACTCACTATAGGGCG
            Class: BAC ends
            High quality sequence start: 19
            High quality sequence stop: 763.
FEATURES             Location/Qualifiers
     source          1. .1217
                     /organism="Gallus gallus"
                     /mol_type="genomic DNA"
                     /strain="Red Jungle Fowl"
                     /db_xref="taxon:9031"
                     /clone="TAM32-17E11"
                     /sex="female"
                     /cell_line="UCD001, inbred 256"
                     /clone_lib="TAM32"
                     /note="Vector: pECBAC1; Site_1: EcoRI; Site_2: EcoRI;
                     TAM32 Female Chicken library - for library and clone
                     ordering information: http://www.hbz.tamu.edu"
ORIGIN
```

```
  Query Match             4.5%;  Score 21;  DB 12;  Length 1217;
  Best Local Similarity   100.0%;  Pred. No. 7.3;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGAAAATAAATGGAACACACA 21
              |||||||||||||||||||||
Db        715 TGAAAATAAATGGAACACACA 695


RESULT 7
CL219260
LOCUS       CL219260                245 bp    DNA     linear   GSS 08-JAN-2004
DEFINITION  ZMMBBc0429L12f ZMMBBc Zea mays genomic clone ZMMBBc0429L12 5',
            genomic survey sequence.
ACCESSION   CL219260
VERSION     CL219260.1  GI:40751934
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 245)
  AUTHORS   Bharti,A.K., Young,S., Kavchok,S., Keizer,G., Bronzino,A.C.,
            Zohovetz,V., Fuks,G., Yu,Y., Wing,R. and Messing,J.
  TITLE     Sequencing of the maize genome at PGIR (2003c)
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Bharti,A.K.
            Dr.Joachim Messing's lab
            The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers
            University
            190 Frelinghuysen Road, Piscataway, NJ 08854, USA
            Tel: 732 445 3801
            Fax: 732 445 5735
            Email: bharti@waksman.rutgers.edu
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .245
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /cultivar="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBBc0429L12"
                     /lab_host="E. coli DH10B"
                     /clone_lib="ZMMBBc"
                     /note="Vector: pTARBAC1.3; Site_1: BamHI; Site_2: BamHI"
ORIGIN

  Query Match             4.3%;  Score 20;  DB 13;  Length 245;
  Best Local Similarity   100.0%;  Pred. No. 25;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        248 TCAAACAGAAGAGCTGATGA 267
              ||||||||||||||||||||
Db          2 TCAAACAGAAGAGCTGATGA 21


RESULT 8
DX327753/c
LOCUS       DX327753                611 bp    DNA     linear   GSS 20-JAN-2006
DEFINITION  OR_ABa0248O05.r OR_ABa Oryza ridleyi genomic clone OR_ABa0248O05
            3', genomic survey sequence.
ACCESSION   DX327753
VERSION     DX327753.1  GI:85604539
KEYWORDS    GSS.
SOURCE      Oryza ridleyi
  ORGANISM  Oryza ridleyi
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BEP
            clade; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 611)
  AUTHORS   Kim,H., Collura,K., Wissotski,M., Byrne,M., Stum,D., Smart,D.,
            Rao,K., Luo,M., Jetty,R., Kudrna,D., Muller,C., Soderlund,C. and
            Wing,R.
  TITLE     OMAP (Oryza Map Alignment Project)- Arizona Genomics Institute
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Rod A. Wing
            Arizona Genomics Institute
            University of Arizona
            Forbes Building Room 303, Tucson, AZ 85721-0036, USA
            Tel: 520 626 9595
            Fax: 520 621 1259
            Email: rwing@genome.arizona.edu
            Plate: 0248  row: O  column: 05
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .611
                     /organism="Oryza ridleyi"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:83308"
                     /clone="OR_ABa0248O05"
                     /tissue_type="leaves"
                     /lab_host="DH10B"
                     /clone_lib="OR_ABa"
                     /note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"
ORIGIN

  Query Match             4.3%;  Score 20;  DB 14;  Length 611;
  Best Local Similarity   100.0%;  Pred. No. 26;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        334 TGTTTTCTTCTCAACCCATG 353
              ||||||||||||||||||||
Db        264 TGTTTTCTTCTCAACCCATG 245


RESULT 9
CX687527
LOCUS       CX687527                613 bp    mRNA    linear   EST 19-JAN-2005
DEFINITION  ydd80c08.y2 Sea urchin EST Lib1 Strongylocentrotus purpuratus cDNA
            clone ydd80c08 5', mRNA sequence.
ACCESSION   CX687527
VERSION     CX687527.1  GI:57948974
KEYWORDS    EST.
SOURCE      Strongylocentrotus purpuratus
  ORGANISM  Strongylocentrotus purpuratus
            Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
            Echinoidea; Euechinoidea; Echinacea; Echinoida;
            Strongylocentrotidae; Strongylocentrotus.
REFERENCE   1  (bases 1 to 613)
  AUTHORS   Coffman,J.A., Robertson,A.J., Clifton,S., Pape,D., Hillier,L.,
            Martin,J., Wylie,T., Dante,M., Meyer,R., Theising,B., Bowers,Y.,
            Gibbons,M., Ronko,I., Tsagareishvili,R., Ritter,E., Kennedy,S. and
            Wilson,R.
  TITLE     WashU Sea Urchin EST Project
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Dr. James A. Coffman
            WashU Sea urchin EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            DNA sequencing by: Washington University Genome Sequencing Center
            Seq primer: -28RPpOT
            High quality sequence stop: 613.
FEATURES             Location/Qualifiers
     source          1. .613
                     /organism="Strongylocentrotus purpuratus"
                     /mol_type="mRNA"
                     /db_xref="taxon:7668"
                     /clone="ydd80c08"
                     /lab_host="DH10B"
                     /clone_lib="Sea urchin EST Lib1"
```

/note="Vector: pCMVSport6.1 (Invitrogen); Site_1: Not1; Site_2: Sma1; Arrayed normalized library of full-length cDNAs representing blastula stage transcriptome of the sea urchin Strongylocentrotus purpuratus, cloned into the vector pCMVSport6.1 (Invitrogen)"
ORIGIN

Query Match 4.3%; Score 20; DB 9; Length 613;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 101 ACGTTGACTGAGCGAGAATT 120
||||||||||||||||||||
Db 502 ACGTTGACTGAGCGAGAATT 521

RESULT 10
CX688942
LOCUS CX688942 626 bp mRNA linear EST 19-JAN-2005
DEFINITION yde22f05.y2 Sea urchin EST Lib1 Strongylocentrotus purpuratus cDNA clone yde22f05 5', mRNA sequence.
ACCESSION CX688942
VERSION CX688942.1 GI:57950601
KEYWORDS EST.
SOURCE Strongylocentrotus purpuratus
ORGANISM Strongylocentrotus purpuratus
Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa; Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae; Strongylocentrotus.
REFERENCE 1 (bases 1 to 626)
AUTHORS Coffman,J.A., Robertson,A.J., Clifton,S., Pape,D., Hillier,L., Martin,J., Wylie,T., Dante,M., Meyer,R., Theising,B., Bowers,Y., Gibbons,M., Ronko,I., Tsagareishvili,R., Ritter,E., Kennedy,S. and Wilson,R.
TITLE WashU Sea Urchin EST Project
JOURNAL Unpublished (2004)
COMMENT Contact: Dr. James A. Coffman
WashU Sea urchin EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
DNA sequencing by: Washington University Genome Sequencing Center
Seq primer: -28RPpOT
High quality sequence stop: 626.
FEATURES Location/Qualifiers
source 1. .626
/organism="Strongylocentrotus purpuratus"
/mol_type="mRNA"
/db_xref="taxon:7668"
/clone="yde22f05"
/lab_host="DH10B"
/clone_lib="Sea urchin EST Lib1"
/note="Vector: pCMVSport6.1 (Invitrogen); Site_1: Not1; Site_2: Sma1; Arrayed normalized library of full-length cDNAs representing blastula stage transcriptome of the sea urchin Strongylocentrotus purpuratus, cloned into the vector pCMVSport6.1 (Invitrogen)"
ORIGIN

Query Match 4.3%; Score 20; DB 9; Length 626;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 101 ACGTTGACTGAGCGAGAATT 120
||||||||||||||||||||
Db 478 ACGTTGACTGAGCGAGAATT 497

RESULT 11
DT673917/c
LOCUS DT673917 640 bp mRNA linear EST 09-SEP-2005
DEFINITION s13dLT65C12FM098_505552 Darnel ryegrass, Lolium temulentum L., Floral meristem Lolium temulentum cDNA, mRNA sequence.
ACCESSION DT673917
VERSION DT673917.1 GI:74338096
KEYWORDS EST.
SOURCE Lolium temulentum
ORGANISM Lolium temulentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BEP clade; Pooideae; Poeae; Lolium.
REFERENCE 1 (bases 1 to 640)
AUTHORS Wang,Z.Y., Mian,M.A.R., Zhang,Y., Zhang,J., Cheng,X., Chen,L., Scott,A.D., Harris,A.R., Gonzales,R.A. and May,G.D.
TITLE ESTs from Darnel ryegrass, Lolium temulentum L., Floral meristem
JOURNAL Unpublished (2005)
COMMENT Contact: May GD
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 224 6650
Fax: 580 224 6692
Email: gdmay@noble.org.
FEATURES Location/Qualifiers
source 1. .640
/organism="Lolium temulentum"
/mol_type="mRNA"
/db_xref="taxon:34176"
/tissue_type="Floral meristem"
/dev_stage="Six-week-old"
/clone_lib="Darnel ryegrass, Lolium temulentum L., Floral meristem"
/note="Vector: Lambda Zap; Floral meristems of Darnel ryegrass, Lolium temulentum L. were isolated from greenhouse plants four days after long-day treatment. The plants were initially grown at short-day conditions for six weeks before the long day treatment. mRNAs were isolated from floral meristems and used for cDNA library construction."
ORIGIN

Query Match 4.3%; Score 20; DB 10; Length 640;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 399 TTTTTGAGTTCTTCAACAGA 418
||||||||||||||||||||
Db 302 TTTTTGAGTTCTTCAACAGA 283

RESULT 12
BH557567/c
LOCUS BH557567 679 bp DNA linear GSS 14-DEC-2001
DEFINITION BOGOG92TF BOGO Brassica oleracea genomic clone BOGOG92, genomic survey sequence.
ACCESSION BH557567
VERSION BH557567.1 GI:17809347
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 679)
AUTHORS Ayele,M., Haas,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S., Utterback,T.R., Wortman,J.R., White,O.R. and Town,C.D.
TITLE Whole genome shotgun sequencing of Brassica oleracea and its application to gene discovery and annotation in Arabidopsis
JOURNAL Genome Res. 15 (4), 487-495 (2005)
PUBMED 15805490
COMMENT Other_GSSs: BOGOG92TR
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.

```
            Tel: 301-838-3523
            Fax: 301-838-0208
            Email: cdtown@tigr.org
            DNA is from a doubled haploid provided by Tom Osborn.
            Seq primer: TF
            Class: sheared ends.
FEATURES                Location/Qualifiers
     source             1. .679
                        /organism="Brassica oleracea"
                        /mol_type="genomic DNA"
                        /strain="TO1000DH3"
                        /db_xref="taxon:3712"
                        /clone="BOGOG92"
                        /clone_lib="BOGO"
                        /note="Vector: pHOS1; Site_1: BstXI; 2-3 kb sheared
                        genomic DNA inserted into pHOS1 using BstXI linkers"
ORIGIN

  Query Match             4.3%;  Score 20;  DB 11;  Length 679;
  Best Local Similarity   100.0%;  Pred. No. 26;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        326 GGTAATGATGTTTTCTTCTC 345
              ||||||||||||||||||||
Db        107 GGTAATGATGTTTTCTTCTC 88


RESULT 13
BZ162280/c
LOCUS       BZ162280                785 bp    DNA     linear   GSS 11-OCT-2002
DEFINITION  CH230-276O24.TV CHORI-230 Segment 2 Rattus norvegicus genomic clone
            CH230-276O24, genomic survey sequence.
ACCESSION   BZ162280
VERSION     BZ162280.1  GI:23803252
KEYWORDS    GSS.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 785)
  AUTHORS   Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K.,
            Shvartsbeyn,A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D.,
            Riggs,F., de Jong,P. and Fraser,C.M.
  TITLE     Rat BAC End Sequences from Library CHORI-230 MboI segment
  JOURNAL   Unpublished (1999)
COMMENT     Other_GSSs: CH230-276O24.TJ
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the rat BAC library CHORI-230
            (http://www.chori.org/bacpac/rat230.htm). For BAC library
            availability, please contact Pieter de Jong (pdejong@mail.cho.org).
            Clones may be purchased from BACPAC Resources
            (http://www.chori.org/bacpac/or ering_information.htm). BAC end
            page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
            Plate: 276  row: O  column: 24
            Seq primer: T7
            Class: BAC ends.
FEATURES                Location/Qualifiers
     source             1. .785
                        /organism="Rattus norvegicus"
                        /mol_type="genomic DNA"
                        /strain="BN/SsNHsd/MCW"
                        /db_xref="taxon:10116"
                        /clone="CH230-276O24"
                        /sex="Female"
                        /cell_type="Brain"
                        /clone_lib="CHORI-230 Segment 2"
                        /note="Vector: pTARBAC1.3; Site_1: MboI; Site_2: MboI;
                        CHORI-230 Rat (BN/SsNHsd/MCW) BAC library produced by
                        Pieter de Jong"
ORIGIN

  Query Match             4.3%;  Score 20;  DB 11;  Length 785;
  Best Local Similarity   100.0%;  Pred. No. 26;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         67 GGTTCAGGAGGCAGGGGCAG 86
              ||||||||||||||||||||
Db        206 GGTTCAGGAGGCAGGGGCAG 187


RESULT 14
BU795670
LOCUS       BU795670                896 bp    mRNA    linear   EST 23-OCT-2003
DEFINITION  SJF2DUA01 SJF Schistosoma japonicum cDNA, mRNA sequence.
ACCESSION   BU795670
VERSION     BU795670.1  GI:28352677
KEYWORDS    EST.
SOURCE      Schistosoma japonicum
  ORGANISM  Schistosoma japonicum
            Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
            Strigeidida; Schistosomatoidea; Schistosomatidae; Schistosoma.
REFERENCE   1  (bases 1 to 896)
  AUTHORS   Hu,W., Yan,Q., Shen,D.K., Liu,F., Zhu,Z.D., Song,H.D., Xu,X.R.,
            Wang,Z.J., Rong,Y.P., Zeng,L.C., Wu,J., Zhang,X., Wang,J.J.,
            Xu,X.N., Wang,S.Y., Fu,G., Zhang,X.L., Wang,Z.Q., Brindley,P.J.,
            McManus,D.P., Xue,C.L., Feng,Z., Chen,Z. and Han,Z.G.
  TITLE     Evolutionary and biomedical implications of a Schistosoma japonicum
            complementary DNA resource
  JOURNAL   Nat. Genet. 35 (2), 139-147 (2003)
  PUBMED    12973349
COMMENT     Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanzg@chgc.sh.cn.
FEATURES                Location/Qualifiers
     source             1. .896
                        /organism="Schistosoma japonicum"
                        /mol_type="mRNA"
                        /db_xref="taxon:6182"
                        /sex="female"
                        /tissue_type="Whole body"
                        /dev_stage="adult"
                        /lab_host="rabbits"
                        /clone_lib="SJF"
ORIGIN

  Query Match             4.3%;  Score 20;  DB 3;  Length 896;
  Best Local Similarity   100.0%;  Pred. No. 26;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 AAAATAAATGGAACACACAT 22
              ||||||||||||||||||||
Db        704 AAAATAAATGGAACACACAT 723


RESULT 15
BB212456/c
LOCUS       BB212456                230 bp    mRNA    linear   EST 30-JUN-2000
DEFINITION  BB212456 RIKEN full-length enriched, 0 day neonate thymus Mus
            musculus cDNA clone A430110O09 3', mRNA sequence.
ACCESSION   BB212456
VERSION     BB212456.1  GI:8877409
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 230)
AUTHORS Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N., Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y., Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y., Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tominaga,N., Toya,T., Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yamanaka,I., Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE RIKEN Mouse ESTs (Konno,H., et al.)
JOURNAL Unpublished (2000)
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Thermostabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J., Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
Please visit our web site (http://genome.rtc.riken.go.jp) for further details.
FEATURES Location/Qualifiers
source 1. .230
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="A430110009"
/tissue_type="thymus"
/dev_stage="0 day neonate"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 0 day neonate thymus"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."
ORIGIN

Query Match 4.1%; Score 19; DB 7; Length 230;
Best Local Similarity 100.0%; Pred. No. 90;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 62 CTGATGGTTCAGGAGGCAG 80
||||||||||||||||||||
Db 39 CTGATGGTTCAGGAGGCAG 21

Search completed: November 4, 2006, 01:08:56
Job time : 2064.56 secs

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 3, 2006, 21:32:07 ; Search time 97.5198 Seconds
(without alignments)
9311.708 Million cell updates/sec

Title: US-10-764-316-5
Perfect score: 468
Sequence: 1 tgaaaataaatggaacacac.....tggatgattgcaggagattt 468

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2461376 seqs, 970166171 residues

Word size : 1

Total number of hits satisfying chosen parameters: 4922276

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published_Applications_NA_New:*
1: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09_NEW_PUB.seq:*
2: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US06_NEW_PUB.seq:*
3: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US07_NEW_PUB.seq:*
4: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US08_NEW_PUB.seq:*
5: /EMC_Celerra_SIDS3/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
6: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10_NEW_PUB.seq:*
7: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11_NEW_PUB.seq:*
8: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11_NEW_PUB.seq1:*
9: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11_NEW_PUB.seq2:*
10: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	468	100.0	468	6	US-10-764-316-5	Sequence 5, Appli
	2	467	99.8	2305	6	US-10-764-316-3	Sequence 3, Appli
	3	300	64.1	412	8	US-11-136-524-10	Sequence 10, Appl
c	4	19	4.1	1458	9	US-11-056-355B-72462	Sequence 72462, A
c	5	19	4.1	1461	9	US-11-056-355B-72470	Sequence 72470, A
c	6	18	3.8	575	8	US-11-266-748A-61868	Sequence 61868, A
c	7	18	3.8	691	6	US-10-953-349-3167	Sequence 3167, Ap
	8	18	3.8	1000	8	US-11-266-748A-114823	Sequence 114823,
c	9	18	3.8	1000	8	US-11-266-748A-119954	Sequence 119954,
	10	18	3.8	1000	8	US-11-266-748A-195502	Sequence 195502,
	11	18	3.8	1000	8	US-11-266-748A-281826	Sequence 281826,
c	12	18	3.8	1000	8	US-11-266-748A-308466	Sequence 308466,
	13	18	3.8	1000	8	US-11-266-748A-390944	Sequence 390944,
c	14	18	3.8	1000	8	US-11-266-748A-481662	Sequence 481662,
	15	18	3.8	1037	8	US-11-266-748A-93124	Sequence 93124, A
c	16	18	3.8	1037	8	US-11-266-748A-145935	Sequence 145935,
	17	18	3.8	1356	9	US-11-348-413-13626	Sequence 13626, A
	18	18	3.8	1952	6	US-10-449-902-28210	Sequence 28210, A
c	19	18	3.8	5526	6	US-10-540-898-796	Sequence 796, App
	20	18	3.8	12610	6	US-10-517-441-441	Sequence 441, App
	21	18	3.8	12610	6	US-10-517-441-715	Sequence 715, App
c	22	18	3.8	84675	6	US-10-539-228-398	Sequence 398, App
c	23	18	3.8	109661	6	US-10-540-898-795	Sequence 795, App
c	24	18	3.8	130877	6	US-10-539-228-54	Sequence 54, Appl
c	25	17	3.6	201	7	US-11-073-360-5989	Sequence 5989, Ap
c	26	17	3.6	484	8	US-11-266-748A-37795	Sequence 37795, A
	27	17	3.6	558	8	US-11-266-748A-162781	Sequence 162781,
	28	17	3.6	570	8	US-11-266-748A-171938	Sequence 171938,
c	29	17	3.6	570	8	US-11-266-748A-245466	Sequence 245466,
	30	17	3.6	598	6	US-10-513-369-3617	Sequence 3617, Ap
	31	17	3.6	604	8	US-11-266-748A-299723	Sequence 299723,
	32	17	3.6	652	8	US-11-266-748A-11817	Sequence 11817, A
	33	17	3.6	686	8	US-11-266-748A-480363	Sequence 480363,
	34	17	3.6	702	8	US-11-266-748A-191891	Sequence 191891,
	35	17	3.6	702	8	US-11-266-748A-226061	Sequence 226061,
c	36	17	3.6	740	9	US-11-056-355B-56754	Sequence 56754, A
	37	17	3.6	807	8	US-11-246-999-22	Sequence 22, Appl
	38	17	3.6	846	8	US-11-266-748A-250972	Sequence 250972,
c	39	17	3.6	846	8	US-11-266-748A-311489	Sequence 311489,
	40	17	3.6	950	8	US-11-266-748A-348672	Sequence 348672,
c	41	17	3.6	950	8	US-11-266-748A-432051	Sequence 432051,
	42	17	3.6	1000	8	US-11-266-748A-198066	Sequence 198066,
c	43	17	3.6	1000	8	US-11-266-748A-205770	Sequence 205770,
	44	17	3.6	1000	8	US-11-266-748A-286159	Sequence 286159,
c	45	17	3.6	1000	8	US-11-266-748A-337588	Sequence 337588,

ALIGNMENTS

RESULT 1
US-10-764-316-5
; Sequence 5, Application US/10764316
; Publication No. US20060127359A1
; GENERAL INFORMATION:
; APPLICANT: BORRELLI, MICHAEL J.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR HEAT ACTIVATED GENE
; TITLE OF INVENTION: THERAPY USING CYTOLETHAL DISTENDING TOXIN
; FILE REFERENCE: 10546-109
; CURRENT APPLICATION NUMBER: US/10/764,316
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: 60/442,473
; PRIOR FILING DATE: 2003-01-24
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 5
; LENGTH: 468
; TYPE: DNA
; ORGANISM: Escherichia coli
US-10-764-316-5

Query Match 100.0%; Score 468; DB 6; Length 468;
Best Local Similarity 100.0%; Pred. No. 4.4e-254;
Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TGAAAATAAATGGAACACACATGTCCGACAACTTGTTACGGGAAGTGGTGCTGTTGATAT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGAAAATAAATGGAACACACATGTCCGACAACTTGTTACGGGAAGTGGTGCTGTTGATAT 60

Qy         61 CCTGATGGTTCAGGAGGCAGGGGCAGTACCAGCTTCTGCAACGTTGACTGAGCGAGAATT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCTGATGGTTCAGGAGGCAGGGGCAGTACCAGCTTCTGCAACGTTGACTGAGCGAGAATT 120

Qy        121 TAGCACTCCTGGTATTCCGATGAATGAGTATATCTGGAATACCGGAACCAATAGTCGCCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TAGCACTCCTGGTATTCCGATGAATGAGTATATCTGGAATACCGGAACCAATAGTCGCCC 180

Qy        181 ACAGGAGTTGTTTATATATTTCTCACGTGTTGATGCATTCGCTAACAGAGTAAATCTTGC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACAGGAGTTGTTTATATATTTCTCACGTGTTGATGCATTCGCTAACAGAGTAAATCTTGC 240

Qy        241 GATTGTTTCAAACAGAAGAGCTGATGAGGTGATTGTATTACCTCCTCCAACTGTTGTATC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GATTGTTTCAAACAGAAGAGCTGATGAGGTGATTGTATTACCTCCTCCAACTGTTGTATC 300

Qy        301 ACGACCGATCATCGGCATTAGAATTGGTAATGATGTTTTCTTCTCAACCCATGCATTGGC 360
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACGACCGATCATCGGCATTAGAATTGGTAATGATGTTTTCTTCTCAACCCATGCATTGGC 360

Qy        361 GAATCGGGGCGTGGATTCAGGAGCAATTGTAAATAGTGTTTTTGAGTTCTTCAACAGACA 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAATCGGGGCGTGGATTCAGGAGCAATTGTAAATAGTGTTTTTGAGTTCTTCAACAGACA 420

Qy        421 AACGGATCCTATAAGACAGGCCGCTAACTGGATGATTGCAGGAGATTT 468
                |||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AACGGATCCTATAAGACAGGCCGCTAACTGGATGATTGCAGGAGATTT 468


RESULT 2
US-10-764-316-3
; Sequence 3, Application US/10764316
; Publication No. US20060127359A1
; GENERAL INFORMATION:
;  APPLICANT: BORRELLI, MICHAEL J.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR HEAT ACTIVATED GENE
;  TITLE OF INVENTION:  THERAPY USING CYTOLETHAL DISTENDING TOXIN
;  FILE REFERENCE: 10546-109
;  CURRENT APPLICATION NUMBER: US/10/764,316
;  CURRENT FILING DATE:  2004-01-23
;  PRIOR APPLICATION NUMBER: 60/442,473
;  PRIOR FILING DATE: 2003-01-24
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 3
;   LENGTH: 2305
;   TYPE: DNA
;   ORGANISM: Escherichia coli
US-10-764-316-3

  Query Match             99.8%;  Score 467;  DB 6;  Length 2305;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-253;
  Matches  467;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 GAAAATAAATGGAACACACATGTCCGACAACTTGTTACGGGAAGTGGTGCTGTTGATATC 61
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1019 GAAAATAAATGGAACACACATGTCCGACAACTTGTTACGGGAAGTGGTGCTGTTGATATC 1078

Qy         62 CTGATGGTTCAGGAGGCAGGGGCAGTACCAGCTTCTGCAACGTTGACTGAGCGAGAATTT 121
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1079 CTGATGGTTCAGGAGGCAGGGGCAGTACCAGCTTCTGCAACGTTGACTGAGCGAGAATTT 1138

Qy        122 AGCACTCCTGGTATTCCGATGAATGAGTATATCTGGAATACCGGAACCAATAGTCGCCCA 181
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1139 AGCACTCCTGGTATTCCGATGAATGAGTATATCTGGAATACCGGAACCAATAGTCGCCCA 1198

Qy        182 CAGGAGTTGTTTATATATTTCTCACGTGTTGATGCATTCGCTAACAGAGTAAATCTTGCG 241
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1199 CAGGAGTTGTTTATATATTTCTCACGTGTTGATGCATTCGCTAACAGAGTAAATCTTGCG 1258

Qy        242 ATTGTTTCAAACAGAAGAGCTGATGAGGTGATTGTATTACCTCCTCCAACTGTTGTATCA 301
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1259 ATTGTTTCAAACAGAAGAGCTGATGAGGTGATTGTATTACCTCCTCCAACTGTTGTATCA 1318

Qy        302 CGACCGATCATCGGCATTAGAATTGGTAATGATGTTTTCTTCTCAACCCATGCATTGGCG 361
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1319 CGACCGATCATCGGCATTAGAATTGGTAATGATGTTTTCTTCTCAACCCATGCATTGGCG 1378

Qy        362 AATCGGGGCGTGGATTCAGGAGCAATTGTAAATAGTGTTTTTGAGTTCTTCAACAGACAA 421
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1379 AATCGGGGCGTGGATTCAGGAGCAATTGTAAATAGTGTTTTTGAGTTCTTCAACAGACAA 1438

Qy        422 ACGGATCCTATAAGACAGGCCGCTAACTGGATGATTGCAGGAGATTT 468
                |||||||||||||||||||||||||||||||||||||||||||||||||
Db       1439 ACGGATCCTATAAGACAGGCCGCTAACTGGATGATTGCAGGAGATTT 1485


RESULT 3
US-11-136-524-10
; Sequence 10, Application US/11136524
; Publication No. US20060094034A1
; GENERAL INFORMATION:
;  APPLICANT: BROUSSEAU, Roland
;  APPLICANT:  HAREL, Josee
;  APPLICANT:  BEKAL, Sadjia
;  TITLE OF INVENTION: ARRAY AND USES THEREOF
;  FILE REFERENCE: 86369-3
;  CURRENT APPLICATION NUMBER: US/11/136,524
;  CURRENT FILING DATE:  2005-05-25
;  NUMBER OF SEQ ID NOS: 182
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10
;   LENGTH: 412
;   TYPE: DNA
;   ORGANISM: Escherichia coli
US-11-136-524-10

  Query Match             64.1%;  Score 300;  DB 8;  Length 412;
  Best Local Similarity   100.0%;  Pred. No. 4.3e-159;
  Matches  300;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        169 CAATAGTCGCCCACAGGAGTTGTTTATATATTTCTCACGTGTTGATGCATTCGCTAACAG 228
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAATAGTCGCCCACAGGAGTTGTTTATATATTTCTCACGTGTTGATGCATTCGCTAACAG 60

Qy        229 AGTAAATCTTGCGATTGTTTCAAACAGAAGAGCTGATGAGGTGATTGTATTACCTCCTCC 288
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGTAAATCTTGCGATTGTTTCAAACAGAAGAGCTGATGAGGTGATTGTATTACCTCCTCC 120

Qy        289 AACTGTTGTATCACGACCGATCATCGGCATTAGAATTGGTAATGATGTTTTCTTCTCAAC 348
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AACTGTTGTATCACGACCGATCATCGGCATTAGAATTGGTAATGATGTTTTCTTCTCAAC 180

Qy        349 CCATGCATTGGCGAATCGGGGCGTGGATTCAGGAGCAATTGTAAATAGTGTTTTTGAGTT 408
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCATGCATTGGCGAATCGGGGCGTGGATTCAGGAGCAATTGTAAATAGTGTTTTTGAGTT 240

Qy        409 CTTCAACAGACAAACGGATCCTATAAGACAGGCCGCTAACTGGATGATTGCAGGAGATTT 468
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTTCAACAGACAAACGGATCCTATAAGACAGGCCGCTAACTGGATGATTGCAGGAGATTT 300


RESULT 4
US-11-056-355B-72462/c
; Sequence 72462, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
;  APPLICANT: Brover, Vyacheslav
;  APPLICANT:  Alexandrov, Nickolai
;  TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
;  TITLE OF INVENTION:  Polypeptides Encoded Thereby
;  FILE REFERENCE: 2750-1590PUS2
;  CURRENT APPLICATION NUMBER: US/11/056,355B
;  CURRENT FILING DATE:  2005-02-14
;  PRIOR APPLICATION NUMBER: 60/544,190
;  PRIOR FILING DATE: 2004-02-13
;  NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 72462
;   LENGTH: 1458
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1458)
;   OTHER INFORMATION: Ceres Seq. ID no. 4935487
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1458)
;   OTHER INFORMATION: Ortholog of  Ceres SEQ ID NO 13574398
;   OTHER INFORMATION: as cited in SEQ ID NO 0
US-11-056-355B-72462

  Query Match             4.1%;  Score 19;  DB 9;  Length 1458;
```

Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        243 TTGTTTCAAACAGAAGAGC 261
              ||||||||||||||||||| 
Db       1281 TTGTTTCAAACAGAAGAGC 1263
```

RESULT 5
US-11-056-355B-72470/c
; Sequence 72470, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nickolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; TITLE OF INVENTION: Polypeptides Encoded Thereby
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 72470
; LENGTH: 1461
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1461)
; OTHER INFORMATION: Ceres Seq. ID no. 4935571
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1461)
; OTHER INFORMATION: Ortholog of Ceres SEQ ID NO 13574398
; OTHER INFORMATION: as cited in SEQ ID NO 0
US-11-056-355B-72470

Query Match 4.1%; Score 19; DB 9; Length 1461;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        243 TTGTTTCAAACAGAAGAGC 261
              ||||||||||||||||||| 
Db       1284 TTGTTTCAAACAGAAGAGC 1266
```

RESULT 6
US-11-266-748A-61868/c
; Sequence 61868, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 61868
; LENGTH: 575
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (438)..(438)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (441)..(441)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (462)..(462)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (476)..(476)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (551)..(551)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (558)..(558)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (565)..(565)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (569)..(569)
; OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-61868

Query Match 3.8%; Score 18; DB 8; Length 575;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        331 TGATGTTTTCTTCTCAAC 348
              ||||||||||||||||||
Db         90 TGATGTTTTCTTCTCAAC 73
```

RESULT 7
US-10-953-349-3167/c
; Sequence 3167, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCONDED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3167
; LENGTH: 691
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-3167

Query Match 3.8%; Score 18; DB 6; Length 691;
Best Local Similarity 100.0%; Pred. No. 12;

```
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        190 GTTTATATATTTCTCACG 207
              ||||||||||||||||||
Db        689 GTTTATATATTTCTCACG 672


RESULT 8
US-11-266-748A-114823
; Sequence 114823, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 114823
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-114823

  Query Match             3.8%;  Score 18;  DB 8;  Length 1000;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        327 GTAATGATGTTTTCTTCT 344
              ||||||||||||||||||
Db        870 GTAATGATGTTTTCTTCT 887


RESULT 9
US-11-266-748A-119954/c
; Sequence 119954, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 119954
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-119954

  Query Match             3.8%;  Score 18;  DB 8;  Length 1000;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        327 GTAATGATGTTTTCTTCT 344
              ||||||||||||||||||
Db        131 GTAATGATGTTTTCTTCT 114


RESULT 10
US-11-266-748A-195502
; Sequence 195502, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 195502
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-195502

  Query Match             3.8%;  Score 18;  DB 8;  Length 1000;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        327 GTAATGATGTTTTCTTCT 344
              ||||||||||||||||||
Db        870 GTAATGATGTTTTCTTCT 887


RESULT 11
```

```
US-11-266-748A-281826
; Sequence 281826, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 281826
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-281826

  Query Match             3.8%;  Score 18;  DB 8;  Length 1000;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        327 GTAATGATGTTTTCTTCT 344
              ||||||||||||||||||
Db        870 GTAATGATGTTTTCTTCT 887


RESULT 12
US-11-266-748A-308466/c
; Sequence 308466, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 308466
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-308466

  Query Match             3.8%;  Score 18;  DB 8;  Length 1000;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        327 GTAATGATGTTTTCTTCT 344
              ||||||||||||||||||
Db        131 GTAATGATGTTTTCTTCT 114


RESULT 13
US-11-266-748A-390944
; Sequence 390944, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 390944
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-390944

  Query Match             3.8%;  Score 18;  DB 8;  Length 1000;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        327 GTAATGATGTTTTCTTCT 344
              ||||||||||||||||||
Db        870 GTAATGATGTTTTCTTCT 887


RESULT 14
US-11-266-748A-481662/c
; Sequence 481662, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
```

; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 481662
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-481662

Query Match 3.8%; Score 18; DB 8; Length 1000;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 327 GTAATGATGTTTTCTTCT 344
||||||||||||||||||
Db 131 GTAATGATGTTTTCTTCT 114

RESULT 15
US-11-266-748A-93124
; Sequence 93124, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 93124
; LENGTH: 1037
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-93124

Query Match 3.8%; Score 18; DB 8; Length 1037;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 405 AGTTCTTCAACAGACAAA 422
||||||||||||||||||
Db 365 AGTTCTTCAACAGACAAA 382

Search completed: November 4, 2006, 01:19:20
Job time : 99.5198 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 3, 2006, 21:28:41 ; Search time 122.65 Seconds
(without alignments)
7139.630 Million cell updates/sec

Title: US-10-764-316-5
Perfect score: 468
Sequence: 1 tgaaaataaatggaacacac.....tggatgattgcaggagattt 468

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1403666 seqs, 935554401 residues

Word size : 1

Total number of hits satisfying chosen parameters: 2806514

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /EMC_Celerra_SIDS3/ptodata/2/ina/1_COMB.seq:*
2: /EMC_Celerra_SIDS3/ptodata/2/ina/5_COMB.seq:*
3: /EMC_Celerra_SIDS3/ptodata/2/ina/6A_COMB.seq:*
4: /EMC_Celerra_SIDS3/ptodata/2/ina/6B_COMB.seq:*
5: /EMC_Celerra_SIDS3/ptodata/2/ina/7_COMB.seq:*
6: /EMC_Celerra_SIDS3/ptodata/2/ina/H_COMB.seq:*
7: /EMC_Celerra_SIDS3/ptodata/2/ina/PCTUS_COMB.seq:*
8: /EMC_Celerra_SIDS3/ptodata/2/ina/PP_COMB.seq:*
9: /EMC_Celerra_SIDS3/ptodata/2/ina/RE_COMB.seq:*
10: /EMC_Celerra_SIDS3/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	18	3.8	1017	3	US-09-328-352-597	Sequence 597, App
2	18	3.8	1356	3	US-09-134-001C-617	Sequence 617, App
3	18	3.8	1356	3	US-09-710-279-2907	Sequence 2907, Ap
c 4	18	3.8	3069	3	US-09-710-279-4352	Sequence 4352, Ap
5	18	3.8	3206	3	US-09-710-279-4382	Sequence 4382, Ap
c 6	18	3.8	4249	3	US-09-710-279-4133	Sequence 4133, Ap
7	18	3.8	34047	3	US-09-949-016-13503	Sequence 13503, A
8	18	3.8	67581	3	US-09-949-016-14768	Sequence 14768, A
9	18	3.8	67643	3	US-09-949-016-14760	Sequence 14760, A
c 10	18	3.8	117838	3	US-09-949-016-17595	Sequence 17595, A
c 11	18	3.8	155617	3	US-09-949-016-16191	Sequence 16191, A
12	18	3.8	193303	3	US-09-497-855A-37	Sequence 37, Appl
13	18	3.8	193303	3	US-09-497-855A-44	Sequence 44, Appl
14	18	3.8	202111	3	US-09-949-016-13877	Sequence 13877, A
15	17	3.6	38	3	US-09-371-772B-11741	Sequence 11741, A
16	17	3.6	38	5	US-10-138-674B-14018	Sequence 14018, A
17	17	3.6	198	3	US-09-248-796A-13245	Sequence 13245, A
18	17	3.6	338	3	US-09-513-999C-33496	Sequence 33496, A
19	17	3.6	601	3	US-09-949-016-42562	Sequence 42562, A
20	17	3.6	601	3	US-09-949-016-42563	Sequence 42563, A
21	17	3.6	601	3	US-09-949-016-42564	Sequence 42564, A
c 22	17	3.6	601	3	US-09-949-016-140295	Sequence 140295,
c 23	17	3.6	601	3	US-09-949-016-140296	Sequence 140296,
c 24	17	3.6	601	3	US-09-949-016-140297	Sequence 140297,
25	17	3.6	601	3	US-09-949-016-160534	Sequence 160534,
26	17	3.6	601	3	US-09-949-016-160535	Sequence 160535,
27	17	3.6	601	3	US-09-949-016-189929	Sequence 189929,
c 28	17	3.6	609	3	US-09-328-111-609	Sequence 609, App
c 29	17	3.6	1164	3	US-09-248-796A-2935	Sequence 2935, Ap
30	17	3.6	1183	3	US-09-949-016-3675	Sequence 3675, Ap
c 31	17	3.6	1197	3	US-09-328-352-1131	Sequence 1131, Ap
c 32	17	3.6	1386	3	US-09-328-352-981	Sequence 981, App
33	17	3.6	1467	3	US-09-489-039A-1303	Sequence 1303, Ap
34	17	3.6	2212	3	US-09-917-254-24	Sequence 24, Appl
35	17	3.6	2237	2	US-08-463-620-1	Sequence 1, Appli
36	17	3.6	2237	2	US-08-224-917-1	Sequence 1, Appli
37	17	3.6	2237	2	US-08-914-853-1	Sequence 1, Appli
38	17	3.6	2237	7	PCT-US95-03934A-1	Sequence 1, Appli
c 39	17	3.6	2412	4	US-10-094-749-241	Sequence 241, App
40	17	3.6	4743	3	US-09-902-540-7191	Sequence 7191, Ap
c 41	17	3.6	5822	3	US-09-902-540-668	Sequence 668, App
42	17	3.6	10357	3	US-08-961-527-191	Sequence 191, App
43	17	3.6	27284	3	US-09-949-016-15417	Sequence 15417, A
44	17	3.6	28720	3	US-09-341-587-7	Sequence 7, Appli
c 45	17	3.6	38343	3	US-09-949-016-15714	Sequence 15714, A

ALIGNMENTS

RESULT 1
US-09-328-352-597
; Sequence 597, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 597
; LENGTH: 1017
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-597

Query Match 3.8%; Score 18; DB 3; Length 1017;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        316 CATTAGAATTGGTAATGA 333
              ||||||||||||||||||
Db        471 CATTAGAATTGGTAATGA 488
```

RESULT 2
US-09-134-001C-617
; Sequence 617, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 617
; LENGTH: 1356
; TYPE: DNA

; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-617

Query Match 3.8%; Score 18; DB 3; Length 1356;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 451 GATGATTGCAGGAGATTT 468
||||||||||||||||||
Db 702 GATGATTGCAGGAGATTT 719

RESULT 3
US-09-710-279-2907
; Sequence 2907, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2907
; LENGTH: 1356
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-710-279-2907

Query Match 3.8%; Score 18; DB 3; Length 1356;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 451 GATGATTGCAGGAGATTT 468
||||||||||||||||||
Db 702 GATGATTGCAGGAGATTT 719

RESULT 4
US-09-710-279-4352/c
; Sequence 4352, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4352
; LENGTH: 3069
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-710-279-4352

Query Match 3.8%; Score 18; DB 3; Length 3069;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 451 GATGATTGCAGGAGATTT 468
||||||||||||||||||
Db 1482 GATGATTGCAGGAGATTT 1465

RESULT 5
US-09-710-279-4382
; Sequence 4382, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4382
; LENGTH: 3206
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-710-279-4382

Query Match 3.8%; Score 18; DB 3; Length 3206;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 451 GATGATTGCAGGAGATTT 468
||||||||||||||||||
Db 2000 GATGATTGCAGGAGATTT 2017

RESULT 6
US-09-710-279-4133/c
; Sequence 4133, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4133
; LENGTH: 4249
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-710-279-4133

Query Match 3.8%; Score 18; DB 3; Length 4249;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 451 GATGATTGCAGGAGATTT 468
||||||||||||||||||
Db 4145 GATGATTGCAGGAGATTT 4128

RESULT 7
US-09-949-016-13503
; Sequence 13503, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.

```
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13503
;   LENGTH: 34047
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(34047)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13503

  Query Match             3.8%;  Score 18;  DB 3;  Length 34047;
  Best Local Similarity   100.0%;  Pred. No. 22;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        329 AATGATGTTTTCTTCTCA 346
              ||||||||||||||||||
Db      20376 AATGATGTTTTCTTCTCA 20393


RESULT 8
US-09-949-016-14768
; Sequence 14768, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14768
;   LENGTH: 67581
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(67581)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14768

  Query Match             3.8%;  Score 18;  DB 3;  Length 67581;
  Best Local Similarity   100.0%;  Pred. No. 22;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        321 GAATTGGTAATGATGTTT 338
              ||||||||||||||||||
Db      10604 GAATTGGTAATGATGTTT 10621


RESULT 9
US-09-949-016-14760
; Sequence 14760, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14760
;   LENGTH: 67643
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(67643)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14760

  Query Match             3.8%;  Score 18;  DB 3;  Length 67643;
  Best Local Similarity   100.0%;  Pred. No. 22;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        321 GAATTGGTAATGATGTTT 338
              ||||||||||||||||||
Db      10592 GAATTGGTAATGATGTTT 10609


RESULT 10
US-09-949-016-17595/c
; Sequence 17595, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17595
;   LENGTH: 117838
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(117838)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17595

  Query Match             3.8%;  Score 18;  DB 3;  Length 117838;
  Best Local Similarity   100.0%;  Pred. No. 22;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        405 AGTTCTTCAACAGACAAA 422
              ||||||||||||||||||
Db      18984 AGTTCTTCAACAGACAAA 18967
```

```
RESULT 11
US-09-949-016-16191/c
; Sequence 16191, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16191
;   LENGTH: 155617
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-16191

  Query Match             3.8%;  Score 18;  DB 3;  Length 155617;
  Best Local Similarity   100.0%;  Pred. No. 22;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        183 AGGAGTTGTTTATATATT 200
              ||||||||||||||||||
Db     108176 AGGAGTTGTTTATATATT 108159


RESULT 12
US-09-497-855A-37
; Sequence 37, Application US/09497855A
; Patent No. 6605432
; GENERAL INFORMATION:
;  APPLICANT: Huang, Tim
;  TITLE OF INVENTION: HIGH-THROUGHPUT METHODS FOR DETECTING DNA METHYLATION
;  FILE REFERENCE: UMO1523
;  CURRENT APPLICATION NUMBER: US/09/497,855A
;  CURRENT FILING DATE:  2000-02-04
;  PRIOR APPLICATION NUMBER: 60/120,592
;  PRIOR FILING DATE: 1999-02-18
;  PRIOR APPLICATION NUMBER: 60/118,760
;  PRIOR FILING DATE: 1999-02-05
;  NUMBER OF SEQ ID NOS: 54
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 37
;   LENGTH: 193303
;   TYPE: DNA
;   ORGANISM: Homo sapiens;
US-09-497-855A-37

  Query Match             3.8%;  Score 18;  DB 3;  Length 193303;
  Best Local Similarity   100.0%;  Pred. No. 22;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        331 TGATGTTTTCTTCTCAAC 348
              ||||||||||||||||||
Db     161051 TGATGTTTTCTTCTCAAC 161068


RESULT 13
US-09-497-855A-44
; Sequence 44, Application US/09497855A
; Patent No. 6605432
; GENERAL INFORMATION:
;  APPLICANT: Huang, Tim
;  TITLE OF INVENTION: HIGH-THROUGHPUT METHODS FOR DETECTING DNA METHYLATION
;  FILE REFERENCE: UMO1523
;  CURRENT APPLICATION NUMBER: US/09/497,855A
;  CURRENT FILING DATE:  2000-02-04
;  PRIOR APPLICATION NUMBER: 60/120,592
;  PRIOR FILING DATE: 1999-02-18
;  PRIOR APPLICATION NUMBER: 60/118,760
;  PRIOR FILING DATE: 1999-02-05
;  NUMBER OF SEQ ID NOS: 54
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 44
;   LENGTH: 193303
;   TYPE: DNA
;   ORGANISM: Homo sapiens;
US-09-497-855A-44

  Query Match             3.8%;  Score 18;  DB 3;  Length 193303;
  Best Local Similarity   100.0%;  Pred. No. 22;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        331 TGATGTTTTCTTCTCAAC 348
              ||||||||||||||||||
Db     161051 TGATGTTTTCTTCTCAAC 161068


RESULT 14
US-09-949-016-13877
; Sequence 13877, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13877
;   LENGTH: 202111
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(202111)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13877

  Query Match             3.8%;  Score 18;  DB 3;  Length 202111;
  Best Local Similarity   100.0%;  Pred. No. 22;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        185 GAGTTGTTTATATATTTC 202
              ||||||||||||||||||
Db      74832 GAGTTGTTTATATATTTC 74849


RESULT 15
US-09-371-772B-11741
; Sequence 11741, Application US/09371772B
; Patent No. 6566127
; GENERAL INFORMATION:
;  APPLICANT: Ribozyme Pharmaceuticals, Inc.
;  APPLICANT:  Pavco, Pam
;  APPLICANT:  McSwiggen, Jim
;  APPLICANT:  Stinchcomb, Dan
;  APPLICANT:  Escobedo, Jaime
;  TITLE OF INVENTION: Method and Reagent for the Treatment of Diseases or Conditions R
;  TITLE OF INVENTION:  Levels of Vascular Endothelial Growth Factor Receptor
```

; FILE REFERENCE: MBHB00,876-J (237/198)
; CURRENT APPLICATION NUMBER: US/09/371,772B
; CURRENT FILING DATE: 1999-08-10
; PRIOR APPLICATION NUMBER: US 60/005,974
; PRIOR FILING DATE: 1995-10-26
; PRIOR APPLICATION NUMBER: US 08/584,040
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 14225
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11741
; LENGTH: 38
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Enzymatic Nucleic Acid
; NAME/KEY: misc_feature
; LOCATION: (31)..(31)
; OTHER INFORMATION: n stands for inosine
US-09-371-772B-11741

Query Match 3.6%; Score 17; DB 3; Length 38;
Best Local Similarity 88.2%; Pred. No. 73;
Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy        253 CAGAAGAGCTGATGAGG 269
              |||||||||:||:||||
Db          1 CAGAAGAGCUGAUGAGG 17
```

Search completed: November 3, 2006, 23:46:17
Job time : 126.65 secs

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 3, 2006, 21:27:32 ; Search time 634.59 Seconds
(without alignments)
9061.940 Million cell updates/sec

Title: US-10-764-316-5
Perfect score: 468
Sequence: 1 tgaaaataaatggaacacac.....tggatgattgcaggagattt 468

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 18892170 seqs, 6143817638 residues

Word size : 1

Total number of hits satisfying chosen parameters: 37781012

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published_Applications_NA_Main:*
1: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US08_PUBCOMB.seq:*
3: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
4: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
5: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
6: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
7: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
8: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
9: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
10: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
11: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
12: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
13: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
14: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11B_PUBCOMB.seq:*
15: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11C_PUBCOMB.seq:*
16: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11D_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	300	64.1	412	9	US-10-425-821-10	Sequence 10, Appl
c	2	20	4.3	83710	9	US-10-417-375-27	Sequence 27, Appl
	3	19	4.1	496	4	US-09-925-065A-426352	Sequence 426352,
	4	19	4.1	496	5	US-09-925-065A-426352	Sequence 426352,
	5	19	4.1	497	12	US-10-301-480-489880	Sequence 489880,
	6	19	4.1	497	12	US-10-301-480-1103289	Sequence 1103289,
c	7	19	4.1	506	12	US-10-301-480-59270	Sequence 59270, A
c	8	19	4.1	506	12	US-10-301-480-59271	Sequence 59271, A
c	9	19	4.1	506	12	US-10-301-480-59272	Sequence 59272, A
c	10	19	4.1	506	12	US-10-301-480-672679	Sequence 672679,
c	11	19	4.1	506	12	US-10-301-480-672680	Sequence 672680,
c	12	19	4.1	506	12	US-10-301-480-672681	Sequence 672681,
c	13	19	4.1	520	6	US-10-027-632-208513	Sequence 208513,
c	14	19	4.1	520	6	US-10-027-632-208514	Sequence 208514,
c	15	19	4.1	520	6	US-10-027-632-208515	Sequence 208515,
c	16	19	4.1	520	7	US-10-027-632-208513	Sequence 208513,
c	17	19	4.1	520	7	US-10-027-632-208514	Sequence 208514,
c	18	19	4.1	520	7	US-10-027-632-208515	Sequence 208515,
	19	19	4.1	615	12	US-10-301-480-489881	Sequence 489881,
	20	19	4.1	615	12	US-10-301-480-1103290	Sequence 1103290,
c	21	19	4.1	1458	16	US-11-096-568A-29902	Sequence 29902, A
c	22	19	4.1	1461	16	US-11-096-568A-29906	Sequence 29906, A
	23	19	4.1	32189	3	US-09-764-878-379	Sequence 379, App
	24	19	4.1	32189	6	US-10-079-854-379	Sequence 379, App
	25	19	4.1	32221	3	US-09-764-878-377	Sequence 377, App
	26	19	4.1	32221	6	US-10-079-854-377	Sequence 377, App
	27	18	3.8	25	8	US-10-719-956-450106	Sequence 450106,
	28	18	3.8	25	9	US-10-719-900-226608	Sequence 226608,
	29	18	3.8	25	16	US-11-136-527-324160	Sequence 324160,
c	30	18	3.8	278	3	US-09-864-408A-3985	Sequence 3985, Ap
	31	18	3.8	423	12	US-10-301-480-53166	Sequence 53166, A
	32	18	3.8	423	12	US-10-301-480-666575	Sequence 666575,
c	33	18	3.8	465	9	US-10-357-930-46151	Sequence 46151, A
	34	18	3.8	474	6	US-10-027-632-199207	Sequence 199207,
	35	18	3.8	474	7	US-10-027-632-199207	Sequence 199207,
	36	18	3.8	481	6	US-10-027-632-174541	Sequence 174541,
	37	18	3.8	481	6	US-10-027-632-174542	Sequence 174542,
	38	18	3.8	481	6	US-10-027-632-298848	Sequence 298848,
	39	18	3.8	481	6	US-10-027-632-298849	Sequence 298849,
	40	18	3.8	481	7	US-10-027-632-174541	Sequence 174541,
	41	18	3.8	481	7	US-10-027-632-174542	Sequence 174542,
	42	18	3.8	481	7	US-10-027-632-298848	Sequence 298848,
	43	18	3.8	481	7	US-10-027-632-298849	Sequence 298849,
	44	18	3.8	503	4	US-09-925-065A-220154	Sequence 220154,
	45	18	3.8	503	5	US-09-925-065A-220154	Sequence 220154,

ALIGNMENTS

RESULT 1
US-10-425-821-10
; Sequence 10, Application US/10425821
; Publication No. US20040219530A1
; GENERAL INFORMATION:
; APPLICANT: BROUSSEAU, Roland
; APPLICANT: HAREL, Jos,e
; APPLICANT: BEKAL, Sadjia
; TITLE OF INVENTION: ARRAY AND USES THEREOF
; FILE REFERENCE: 86369-3
; CURRENT APPLICATION NUMBER: US/10/425,821
; CURRENT FILING DATE: 2003-04-30
; NUMBER OF SEQ ID NOS: 176
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10
; LENGTH: 412
; TYPE: DNA
; ORGANISM: Escherichia coli
US-10-425-821-10

Query Match 64.1%; Score 300; DB 9; Length 412;
Best Local Similarity 100.0%; Pred. No. 7.3e-156;
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        169 CAATAGTCGCCCACAGGAGTTGTTTATATATTTCTCACGTGTTGATGCATTCGCTAACAG 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAATAGTCGCCCACAGGAGTTGTTTATATATTTCTCACGTGTTGATGCATTCGCTAACAG 60

Qy        229 AGTAAATCTTGCGATTGTTTCAAACAGAAGAGCTGATGAGGTGATTGTATTACCTCCTCC 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGTAAATCTTGCGATTGTTTCAAACAGAAGAGCTGATGAGGTGATTGTATTACCTCCTCC 120

Qy        289 AACTGTTGTATCACGACCGATCATCGGCATTAGAATTGGTAATGATGTTTTCTTCTCAAC 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AACTGTTGTATCACGACCGATCATCGGCATTAGAATTGGTAATGATGTTTTCTTCTCAAC 180

Qy        349 CCATGCATTGGCGAATCGGGGCGTGGATTCAGGAGCAATTGTAAATAGTGTTTTTGAGTT 408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCATGCATTGGCGAATCGGGGCGTGGATTCAGGAGCAATTGTAAATAGTGTTTTTGAGTT 240
```

```
Qy        409 CTTCAACAGACAAACGGATCCTATAAGACAGGCCGCTAACTGGATGATTGCAGGAGATTT 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTTCAACAGACAAACGGATCCTATAAGACAGGCCGCTAACTGGATGATTGCAGGAGATTT 300
```

RESULT 2
US-10-417-375-27/c
; Sequence 27, Application US/10417375
; Publication No. US20040219528A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Therapeutic Targets in Cancer
; FILE REFERENCE: 529452001600
; CURRENT APPLICATION NUMBER: US/10/417,375
; CURRENT FILING DATE: 2003-04-15
; NUMBER OF SEQ ID NOS: 176
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 83710
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(83710)
; OTHER INFORMATION: n = A,T,C or G
US-10-417-375-27

Query Match 4.3%; Score 20; DB 9; Length 83710;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        118 ATTTAGCACTCCTGGTATTC 137
              ||||||||||||||||||||
Db       3171 ATTTAGCACTCCTGGTATTC 3152
```

RESULT 3
US-09-925-065A-426352
; Sequence 426352, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 426352
; LENGTH: 496
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-426352

Query Match 4.1%; Score 19; DB 4; Length 496;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        326 GGTAATGATGTTTTCTTCT 344
              |||||||||||||||||||
Db        158 GGTAATGATGTTTTCTTCT 176
```

RESULT 4
US-09-925-065A-426352
; Sequence 426352, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 426352
; LENGTH: 496
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-426352

Query Match 4.1%; Score 19; DB 5; Length 496;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        326 GGTAATGATGTTTTCTTCT 344
              |||||||||||||||||||
Db        158 GGTAATGATGTTTTCTTCT 176
```

RESULT 5
US-10-301-480-489880
; Sequence 489880, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 489880
; LENGTH: 497
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-489880

Query Match 4.1%; Score 19; DB 12; Length 497;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        326 GGTAATGATGTTTTCTTCT 344
              |||||||||||||||||||
Db        158 GGTAATGATGTTTTCTTCT 176
```

RESULT 6

```
US-10-301-480-1103289
; Sequence 1103289, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1103289
;   LENGTH: 497
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-1103289

  Query Match             4.1%;  Score 19;  DB 12;  Length 497;
  Best Local Similarity   100.0%;  Pred. No. 27;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        326 GGTAATGATGTTTTCTTCT 344
              |||||||||||||||||||
Db        158 GGTAATGATGTTTTCTTCT 176


RESULT 7
US-10-301-480-59270/c
; Sequence 59270, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 59270
;   LENGTH: 506
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-59270

  Query Match             4.1%;  Score 19;  DB 12;  Length 506;
  Best Local Similarity   100.0%;  Pred. No. 27;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         62 CTGATGGTTCAGGAGGCAG 80
              |||||||||||||||||||
Db        241 CTGATGGTTCAGGAGGCAG 223


RESULT 8
US-10-301-480-59271/c
; Sequence 59271, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 59271
;   LENGTH: 506
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-59271

  Query Match             4.1%;  Score 19;  DB 12;  Length 506;
  Best Local Similarity   100.0%;  Pred. No. 27;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         62 CTGATGGTTCAGGAGGCAG 80
              |||||||||||||||||||
Db        241 CTGATGGTTCAGGAGGCAG 223


RESULT 9
US-10-301-480-59272/c
; Sequence 59272, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 59272
;   LENGTH: 506
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-59272

  Query Match             4.1%;  Score 19;  DB 12;  Length 506;
  Best Local Similarity   100.0%;  Pred. No. 27;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         62 CTGATGGTTCAGGAGGCAG 80
              |||||||||||||||||||
Db        241 CTGATGGTTCAGGAGGCAG 223


RESULT 10
US-10-301-480-672679/c
; Sequence 672679, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 672679
;   LENGTH: 506
```

; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-672679

Query Match 4.1%; Score 19; DB 12; Length 506;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 62 CTGATGGTTCAGGAGGCAG 80
||||||||||||||||||||
Db 241 CTGATGGTTCAGGAGGCAG 223

RESULT 11
US-10-301-480-672680/c
; Sequence 672680, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 672680
; LENGTH: 506
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-672680

Query Match 4.1%; Score 19; DB 12; Length 506;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 62 CTGATGGTTCAGGAGGCAG 80
||||||||||||||||||||
Db 241 CTGATGGTTCAGGAGGCAG 223

RESULT 12
US-10-301-480-672681/c
; Sequence 672681, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 672681
; LENGTH: 506
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-672681

Query Match 4.1%; Score 19; DB 12; Length 506;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 62 CTGATGGTTCAGGAGGCAG 80
||||||||||||||||||||
Db 241 CTGATGGTTCAGGAGGCAG 223

RESULT 13
US-10-027-632-208513/c
; Sequence 208513, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 208513
; LENGTH: 520
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-208513

Query Match 4.1%; Score 19; DB 6; Length 520;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 62 CTGATGGTTCAGGAGGCAG 80
||||||||||||||||||||
Db 240 CTGATGGTTCAGGAGGCAG 222

RESULT 14
US-10-027-632-208514/c
; Sequence 208514, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 208514
; LENGTH: 520
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-208514

Query Match 4.1%; Score 19; DB 6; Length 520;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         62 CTGATGGTTCAGGAGGCAG 80
              ||||||||||||||||||||
Db        240 CTGATGGTTCAGGAGGCAG 222
```

RESULT 15
US-10-027-632-208515/c
; Sequence 208515, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 208515
; LENGTH: 520
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-208515

Query Match 4.1%; Score 19; DB 6; Length 520;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         62 CTGATGGTTCAGGAGGCAG 80
              ||||||||||||||||||||
Db        240 CTGATGGTTCAGGAGGCAG 222
```

Search completed: November 3, 2006, 23:12:50
Job time : 636.59 secs

Page Blank (use)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 3, 2006, 20:16:55 ; Search time 273.593 Seconds
(without alignments)
11926.532 Million cell updates/sec

Title: US-10-764-316-5
Perfect score: 468
Sequence: 1 tgaaaataaatggaacacac.....tggatgattgcaggagattt 468

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 5244920 seqs, 3486124231 residues

Word size : 1

Total number of hits satisfying chosen parameters: 10489196

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_8:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*
15: geneseqn2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	Query Match	Length	DB	ID	Description
1		300	64.1	412	13	ADT98253	Adt98253 E. coli c
2		44	9.4	822	4	AAF85771	Aaf85771 E coli cy
3	c	20	4.3	83102	13	ABD32667	Abd32667 Mouse can
4		19	4.1	32189	5	AAS30115	Aas30115 Human lun
5		19	4.1	32189	10	ADB33452	Adb33452 Human nov
6		19	4.1	32221	5	AAS30113	Aas30113 Human lun
7		19	4.1	32221	10	ADB33450	Adb33450 Human nov
8	c	18	3.8	278	6	ABN77046	Abn77046 Human ORF
9		18	3.8	309	10	ADL24711	Adl24711 Intestina
10		18	3.8	459	13	ACF85026	Acf85026 Human SIR
11	c	18	3.8	465	5	ABV46132	Abv46132 Human pro
12	c	18	3.8	565	3	AAC35341	Aac35341 Arabidops
13	c	18	3.8	565	15	AEF29701	Aef29701 Lead_Cere
14	c	18	3.8	575	4	AAH12181	Aah12181 Human cDN
15		18	3.8	675	3	AAA70256	Aaa70256 Plasmodiu
16		18	3.8	758	4	AAH04264	Aah04264 Human cDN
17	c	18	3.8	826	4	ABL14609	Abl14609 Drosophil
18		18	3.8	1017	9	ADA29310	Ada29310 DNA encod
19		18	3.8	1356	4	AAH53757	Aah53757 S. epider
20		18	3.8	1356	6	ABN91154	Abn91154 Staphyloc
21		18	3.8	1356	13	ADS02625	Ads02625 Staphyloc
22		18	3.8	1409	4	AAK84295	Aak84295 Human imm
23		18	3.8	1410	4	AAK84294	Aak84294 Human imm
24		18	3.8	1944	4	AAH16394	Aah16394 Human cDN
25	c	18	3.8	2868	4	AAH15964	Aah15964 Human cDN
26		18	3.8	2890	4	ABL14608	Abl14608 Drosophil
27	c	18	3.8	3008	13	ADR07926	Adr07926 Full leng
28	c	18	3.8	3069	4	AAH54988	Aah54988 S. epider
29		18	3.8	3206	4	AAH55018	Aah55018 S. epider
30	c	18	3.8	4249	4	AAH54769	Aah54769 S. epider
31	c	18	3.8	5526	12	ADQ97819	Adq97819 Human can
32		18	3.8	7434	6	AAD28387	Aad28387 Human che
33		18	3.8	10324	4	AAK65263	Aak65263 Human imm
34		18	3.8	12610	13	ADS89699	Ads89699 Oligonucl
35		18	3.8	12610	13	ADS89425	Ads89425 Oligonucl
36		18	3.8	14136	4	AAK65264	Aak65264 Human imm
37		18	3.8	14387	4	AAK65265	Aak65265 Human imm
38		18	3.8	27324	4	ABA89226	Aba89226 Escherich
39	c	18	3.8	32249	4	AAI99368	Aai99368 Human exc
40	c	18	3.8	32249	5	AAI63718	Aai63718 Human kid
41		18	3.8	65011	6	ABS79015	Abs79015 E. coli C
42		18	3.8	68611	10	ADH80582	Adh80582 Escherich
43	c	18	3.8	84675	13	ABD33325	Abd33325 Human can
44	c	18	3.8	109661	12	ADQ97818	Adq97818 Human can
45	c	18	3.8	130877	13	ABD33104	Abd33104 Human can

ALIGNMENTS

RESULT 1
ADT98253
ID ADT98253 standard; DNA; 412 BP.
XX
AC ADT98253;
XX
DT 27-JAN-2005 (first entry)
XX
DE E. coli cdt1 DNA.
XX
KW ds; array; pathotype; microorganism detection; infection; probe;
KW virulence gene; enterotoxigenic E. coli; ETEC; enteropathogenic E. coli;
KW EPEC; enterohemorrhagic E. coli; EHEC; enteroaggregative E. coli; EAEC;
KW enteroinvasive E. coli; EIEC; uropathogenic strain; UPEC;
KW neonatal meningitis; MENEC; septicemia; SEPEC; cell-detaching E. coli;
KW CDEC; diffusely adherent E. coli; DAEC; toxin; adhesion factor;
KW secretory system protein; capsule antigen; somatic antigen;
KW flagellar antigen; invasin; autotransporter protein;
KW aerobactin system protein; environment; food; water; air; soil; blood;
KW urine; amniotic fluid; feces; diagnosis; diarrhea; hemorrhagic colitis;
KW hemolytic uremic syndrome; invasive intestinal infection; dysentery;
KW urinary tract infection; neonatal meningitis; septicemia; cdt1.
XX
OS Escherichia coli.
XX
PN US2004219530-A1.
XX
PD 04-NOV-2004.
XX
PF 30-APR-2003; 2003US-00425821.
XX
PR 30-APR-2003; 2003US-00425821.
XX
PA (BROU/) BROUSSEAU R.
PA (HARE/) HAREL J.
PA (BEKA/) BEKAL S.
XX
PI Brousseau R, Harel J, Bekal S;
XX
DR WPI; 2004-774932/76.
XX

```
PT   New array of nucleic acid probes useful for identifying or characterizing
PT   a pathotype of a microorganism, for detecting the presence of a
PT   microorganism (e.g. Escherichia coli) or for assessing infections and
PT   diseases in a subject.
XX
PS   Claim 12; SEQ ID NO 10; 76pp; English.
XX
CC   This invention describes a novel array and methods useful for identifying
CC   and/or characterizing a pathotype of a microorganism, for detecting the
CC   presence of a microorganism (e.g. E. coli) in a sample, or for assessing
CC   infections and diseases in a subject. The method comprises a substrate
CC   and nucleic acid probes, each of the probes being bound to the substrate
CC   at a discrete location. The probes comprise a first probe for a first
CC   pathotype of a species of a microorganism and a second probe for a second
CC   pathotype of the species, where the first and second pathotypes are not
CC   identical. The invention can be used to make a commercial package
CC   comprising the array together with instructions for: (a) detecting the
CC   presence of a microorganism in a sample; (b) determining the pathotype of
CC   a microorganism in a sample; (c) diagnosing an infection by a
CC   microorganism in a subject; (d) diagnosing a condition related to
CC   infection by a microorganism in a subject; or (e) any combination of (a)-
CC   (d). The probe is for a virulence gene or its fragment or a substantially
CC   identical sequence, where the virulence gene is associated with
CC   pathogenicity of the microorganism. The microorganism is a bacterium of
CC   the family Enterobacteriaceae, particularly Escherichia coli. The first
CC   and second pathotypes each independently comprise a pathotype selected
CC   from: (a) enterotoxigenic E. coli (ETEC); (b) enteropathogenic E. coli
CC   (EPEC); (c) enterohemorrhagic E. coli (EHEC); (d) enteroaggregative E.
CC   coli (EAEC); (e) enteroinvasive E. coli (EIEC); (f) uropathogenic strains
CC   (UPEC); (g) E. coli strains involved in neonatal meningitis (MENEC); (h)
CC   E. coli strains involved in septicemia (SEPEC); (i) cell-detaching E.
CC   coli (CDEC); and (j) diffusely adherent E. coli (DAEC). The virulence
CC   gene encodes a polypeptide of a class of proteins selected from toxins,
CC   adhesion factors, secretory system proteins, capsule antigens, somatic
CC   antigens, flagellar antigens, invasins, autotransporter proteins, and
CC   aerobactin system proteins. The virulence gene is selected from afaBC3,
CC   afaE5, afaE7, afaD8, aggA, aggC, aida, bfpA, bmaE, cdt1, cdt2, cdt3,
CC   cfaI, clpG, cnf1, cnf2, cs1, cs3, cs31a, cvaC, derb122, eae, eaf, east1,
CC   ehxA, espA group I, espA group II, espA group III, espB group I, espB
CC   group II, espB group III, espC, espP, etpD, F17A, F17G, F18, F4, F41, F5,
CC   F6, fimA group I, fimA group II, fimH, fliC, focG, fyuA, hlyA, hlyC,
CC   ibe10, iha, invX, ipaC, iroN, irp1, irp2, iss, iucD, iutA, katP, kfiB,
CC   kpsMTII, kpsMTIII, 17095, leoA, IngA, lt, neuC, nfaE, ompA, ompT, paa,
CC   papAH, papC, papEF, papG group I, papG group II, papG group III, pai,
CC   rfb09, rfbO101, rfbO111, rfbE O157, rfbE O157 H7, rfc O4, rtx, sfaDE,
CC   sfaA, stah, stap, stb, stx1, stx2, stxAI, stxAII, stxB I, stx B II, stxB
CC   III, tir group I, tir group II, tir group III, traT, and tsh. The sample
CC   is selected from environmental samples, biological samples and food. The
CC   environmental samples are selected from water, air and soil. The
CC   biological samples are selected from blood, urine, amniotic fluid, feces,
CC   tissues, cells, cell cultures and biological secretions, excretions and
CC   discharge. The method can be used for diagnosing a condition related to
CC   infection by the microorganism in a subject, where the condition is
CC   diarrhea, hemorrhagic colitis, hemolytic uremic syndrome, invasive
CC   intestinal infections, dysentery, urinary tract infections, neonatal
CC   meningitis or septicemia. ADT98244-ADT98347 represent E. coli genes used
CC   as targets in the array of the invention. ADT98348-ADT98419 represent PCR
CC   primers used to amplify the target genes.
XX
SQ   Sequence 412 BP; 112 A; 85 C; 89 G; 126 T; 0 U; 0 Other;

  Query Match             64.1%;  Score 300;  DB 13;  Length 412;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-147;
  Matches  300;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        169 CAATAGTCGCCCACAGGAGTTGTTTATATATTTCTCACGTGTTGATGCATTCGCTAACAG 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAATAGTCGCCCACAGGAGTTGTTTATATATTTCTCACGTGTTGATGCATTCGCTAACAG 60

Qy        229 AGTAAATCTTGCGATTGTTTCAAACAGAAGAGCTGATGAGGTGATTGTATTACCTCCTCC 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGTAAATCTTGCGATTGTTTCAAACAGAAGAGCTGATGAGGTGATTGTATTACCTCCTCC 120

Qy        289 AACTGTTGTATCACGACCGATCATCGGCATTAGAATTGGTAATGATGTTTTCTTCTCAAC 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AACTGTTGTATCACGACCGATCATCGGCATTAGAATTGGTAATGATGTTTTCTTCTCAAC 180

Qy        349 CCATGCATTGGCGAATCGGGGCGTGGATTCAGGAGCAATTGTAAATAGTGTTTTTGAGTT 408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCATGCATTGGCGAATCGGGGCGTGGATTCAGGAGCAATTGTAAATAGTGTTTTTGAGTT 240

Qy        409 CTTCAACAGACAAACGGATCCTATAAGACAGGCCGCTAACTGGATGATTGCAGGAGATTT 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTTCAACAGACAAACGGATCCTATAAGACAGGCCGCTAACTGGATGATTGCAGGAGATTT 300


RESULT 2
AAF85771
ID   AAF85771 standard; DNA; 822 BP.
XX
AC   AAF85771;
XX
DT   10-DEC-2001  (first entry)
XX
DE   E coli cytolethal distending toxin CdtB coding sequence.
XX
KW   Cytolethal distending toxin; CdtA; CdtB; actin cytoskeleton; DNase;
KW   mitotic block; proliferative cell disorder; cancer; psoriasis;
KW   pathogen infection; ds.
XX
OS   Escherichia coli.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .822
FT                   /*tag=  a
FT                   /product= "CdtB"
XX
PN   WO200134205-A2.
XX
PD   17-MAY-2001.
XX
PF   13-NOV-2000; 2000WO-GB004329.
XX
PR   12-NOV-1999;   99GB-00026875.
XX
PA   (MICR-) MICROBIOLOGICAL RES AUTHORITY.
XX
PI   Purdy D,  Charlton S,  Henderson I;
XX
DR   WPI; 2001-328941/34.
DR   P-PSDB; AAB60971.
XX
PT   Agent for treating proliferative disorders and intracellular infection
PT   comprises a targeting moiety and a component.
XX
PS   Claim 19; Page 37; 38pp; English.
XX
CC   The present invention provides an agent comprising a targeting moiety and
CC   an agent capable of inhibiting cell division in a target cell of
CC   interest. The second component may be the cytolethal distending toxin
CC   CdtB, which is encoded by the present sequence. This can be used in the
CC   treatment of cell proliferation disorders, including cancer and
CC   psoriasis, to inhibit mitosis, or to treat infection by an intracellular
CC   pathogen
XX
SQ   Sequence 822 BP; 229 A; 157 C; 188 G; 248 T; 0 U; 0 Other;

  Query Match             9.4%;  Score 44;  DB 4;  Length 822;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-12;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        425 GATCCTATAAGACAGGCCGCTAACTGGATGATTGCAGGAGATTT 468
              ||||||||||||||||||||||||||||||||||||||||||||
Db        532 GATCCTATAAGACAGGCCGCTAACTGGATGATTGCAGGAGATTT 575
```

RESULT 3
ABD32667/c
ID ABD32667 standard; DNA; 83102 BP.
XX
AC ABD32667;
XX
DT 18-NOV-2004 (first entry)
XX
DE Mouse cancer-associated genomic DNA MD13-073.
XX
KW Mouse; ds; cancer-associated protein; gene; cytostatic; cancer;
KW leukaemia; lymphoma; CAP.
XX
OS Mus musculus.
XX
PN WO2004074320-A2.
XX
PD 02-SEP-2004.
XX
PF 17-FEB-2004; 2004WO-US004730.
XX
PR 14-FEB-2003; 2003US-00367094.
PR 14-MAR-2003; 2003US-00388838.
PR 15-APR-2003; 2003US-00417375.
PR 13-JUN-2003; 2003US-00461862.
PR 15-SEP-2003; 2003US-00663431.
PR 15-DEC-2003; 2003US-00737318.
XX
PA (SAGR-) SAGRES DISCOVERY INC.
XX
PI Morris DW, Morris DW, Malandro MS;
XX
DR WPI; 2004-652914/63.
XX
PT New isolated cancer-associated polynucleotides and polypeptides useful
PT for diagnosing, preventing or treating cancers, especially lymphoma and
PT leukemia, or in screening for agents that modulate cancer.
XX
PS disclosure; seqid 212; 310pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising at least 10
CC contiguous nucleotides of any of the 233 polynucleotide sequences given
CC in the specification, or its complement. The nucleic acids encode cancer-
CC associated proteins. Also included are an expression vector comprising
CC the isolated nucleic acid cited above, a host cell comprising the above
CC recombinant nucleic acid or expression vector, a microarray for detecting
CC a cancer-associated (CA) nucleic acid comprising at least one probe
CC comprising at least 10 contiguous nucleotides of any of the above-
CC mentioned nucleotide sequences, an isolated polypeptide (encoded within
CC an open reading frame of a CA sequence selected from any of the 95
CC polynucleotide sequences as mentioned in the specification, or its
CC complement), an isolated antibody, (or its antigen binding fragment) that
CC binds to the above polypeptide, a hybridoma that produces the above
CC monoclonal antibody, a pharmaceutical composition comprising the above
CC antibody and a pharmaceutical excipient, a kit for detecting cancer
CC cells(comprising the antibody cited above, methods for diagnosing cancer
CC or for detecting the presence or absence of cancer cells in an
CC individual, a method for inhibiting growth of cancer cells in an
CC individual, a method for delivering a therapeutic agent to cancer cells
CC in an individual, an electronic library comprising the above
CC polynucleotide or polypeptide (or their fragments), methods of screening
CC for anticancer activity or for a bioactive agent capable of modulating
CC the activity of a CA protein (CAP), methods for detecting cancer
CC associated with expression of a polypeptide in a test cell sample, a
CC method for treating cancers and a method for inhibiting the expression of
CC CA gene in a cell. The composition and methods are useful for detecting,
CC diagnosing, preventing and treating cancers, especially lymphoma and
CC leukaemia. These may also be used in screening for agents that modulate
CC cancer. The present sequence is a mouse CAP genomic sequence. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX

SQ Sequence 83102 BP; 23417 A; 17147 C; 17240 G; 24360 T; 0 U; 938 Other;

Query Match 4.3%; Score 20; DB 13; Length 83102;
Best Local Similarity 100.0%; Pred. No. 7.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 118 ATTTAGCACTCCTGGTATTC 137
||||||||||||||||||||
Db 3009 ATTTAGCACTCCTGGTATTC 2990

RESULT 4
AAS30115
ID AAS30115 standard; DNA; 32189 BP.
XX
AC AAS30115;
XX
DT 21-NOV-2001 (first entry)
XX
DE Human lung antigen genomic DNA #185.
XX
KW Lung antigen protein; human; mouse; rabbit; goat; horse; cat; dog;
KW chicken; sheep; immunosuppressive; antiarthritic; vasotropic;
KW antirheumatic; antiproliferative; cytostatic; cardiant; neuroprotective;
KW cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;
KW ophthalmological; vulnerary; gene therapy; autoimmune disease; neoplasm;
KW hyperproliferative disorder; breast; liver; cardiovascular disorder; ds;
KW cerebrovascular disorder; nervous system disorder; bacterial infection;
KW fungal infection; viral infection; ocular disorder; endocrine disorder;
KW gastrointestinal disorder; renal disorder; respiratory disorder;
KW wound healing; skin aging; organ transplantation; food preservative;
KW tissue regeneration; anti-infertility; food additive.
XX
OS Homo sapiens.
XX
PN WO200155303-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001301.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.

PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-457723/49.
XX
PT Isolated polypeptide for treating, preventing and/ or prognosing
PT respiratory disorders related to the lung including lung cancers and also
PT for testing and detection e.g. diagnosis.
XX
PS Claim 1; SEQ ID NO 379; 507pp; English.
XX
CC Sequences AAS29931-AAS30164 represent genomic DNA molecules, which encode
CC the lung antigen polypeptides of the invention. Lung antigen polypeptides
CC and their associated polynucleotides are useful in the diagnosis,
CC treatment and prevention of various types of disorders in e.g. humans,
CC mice, rabbits, goats, horses, cats, dogs, chickens or sheep. A
CC pathological condition can be determined by detecting the presence or
CC absence of a mutation in a lung antigen polynucleotide. The treatable
CC disorders include autoimmune diseases such as rheumatoid arthritis,
CC hyperproliferative disorders such as neoplasms of the breast or liver,
CC cardiovascular disorders such as cardiac arrest, cerebrovascular
CC disorders such as cerebral ischaemia, nervous system disorders such as
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi,
CC ocular disorders such as corneal infection, endocrine disorders such as
CC premature labour and infertility, gastrointestinal disorders such as
CC Crohn's disease, renal disorders such as glomerulonephritis and
CC respiratory disorders such as asthma and pleurisy. The polypeptides can
CC also be used to aid wound healing, to prevent skin aging due to sunburn,
CC to maintain organs before transplantation, to regenerate tissues and in
CC chemotaxis. The polypeptides can also be used as a food additive or
CC preservative to increase or decrease storage capabilities. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences

Query Match 4.1%; Score 19; DB 5; Length 32189;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 62 CTGATGGTTCAGGAGGCAG 80

```
                    ||||||||||||||||||
Db      23959 CTGATGGTTCAGGAGGCAG 23977

RESULT 5
ADB33452
ID   ADB33452 standard; DNA; 32189 BP.
XX
AC   ADB33452;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Human novel lung related polypeptide DNA SEQ ID NO 379.
XX
KW   gene therapy; lung antigen; neoplasia; acute myelogenous leukaemia;
KW   adenocarcinoma; respiratory disorder; chronic rhinitis; sinusitis;
KW   immunodeficiency; X-linked agammaglobulinaemia;
KW   X-linked infantile agammaglobulinaemia; inflammatory disorder;
KW   adrenalitis; alveolitis; immune complex disease; serum sickness;
KW   polyarteritis nodosa; bleeding disorder; thrombocytopenia;
KW   Von Willebrand's disease; acquired platelet dysfunction; kidney failure;
KW   multiple myeloma; macrophage related disorder; Gaucher's disease;
KW   Neimann-Pick disease; tumour; colon cancer; pancreatic cancer;
KW   renal disorder; nephritis; bone disorder; Albers-Schonberg disease;
KW   bowleg; muscle disorder; Becker's muscular dystrophy;
KW   Duchenne's muscular dystrophy; nervous disorder; ischaemic lesion;
KW   traumatic lesion; endocrine disorder; Cushing's syndrome;
KW   corticosteroid deficiency; gastrointestinal disorder; dysphagia;
KW   gastric reflux; human; ds.
XX
OS   Homo sapiens.
XX
PN   US2003054368-A1.
XX
PD   20-MAR-2003.
XX
PF   22-FEB-2002; 2002US-00079854.
XX
PR   31-JAN-2000; 2000US-0179065P.
PR   04-FEB-2000; 2000US-0180628P.
PR   24-FEB-2000; 2000US-0184664P.
PR   02-MAR-2000; 2000US-0186350P.
PR   16-MAR-2000; 2000US-0189874P.
PR   17-MAR-2000; 2000US-0190076P.
PR   18-APR-2000; 2000US-0198123P.
PR   19-MAY-2000; 2000US-0205515P.
PR   07-JUN-2000; 2000US-0209467P.
PR   28-JUN-2000; 2000US-0214886P.
PR   30-JUN-2000; 2000US-0215135P.
PR   07-JUL-2000; 2000US-0216647P.
PR   07-JUL-2000; 2000US-0216880P.
PR   11-JUL-2000; 2000US-0217487P.
PR   11-JUL-2000; 2000US-0217496P.
PR   14-JUL-2000; 2000US-0218290P.
PR   26-JUL-2000; 2000US-0220963P.
PR   26-JUL-2000; 2000US-0220964P.
PR   14-AUG-2000; 2000US-0224518P.
PR   14-AUG-2000; 2000US-0224519P.
PR   14-AUG-2000; 2000US-0225213P.
PR   14-AUG-2000; 2000US-0225214P.
PR   14-AUG-2000; 2000US-0225266P.
PR   14-AUG-2000; 2000US-0225267P.
PR   14-AUG-2000; 2000US-0225268P.
PR   14-AUG-2000; 2000US-0225270P.
PR   14-AUG-2000; 2000US-0225447P.
PR   14-AUG-2000; 2000US-0225757P.
PR   14-AUG-2000; 2000US-0225758P.
PR   14-AUG-2000; 2000US-0225759P.
PR   18-AUG-2000; 2000US-0226279P.
PR   22-AUG-2000; 2000US-0226681P.
PR   22-AUG-2000; 2000US-0226868P.
PR   22-AUG-2000; 2000US-0227182P.
PR   23-AUG-2000; 2000US-0227009P.
PR   30-AUG-2000; 2000US-0228924P.
PR   01-SEP-2000; 2000US-0229287P.
PR   01-SEP-2000; 2000US-0229343P.
PR   01-SEP-2000; 2000US-0229344P.
PR   01-SEP-2000; 2000US-0229345P.
PR   05-SEP-2000; 2000US-0229509P.
PR   05-SEP-2000; 2000US-0229513P.
PR   06-SEP-2000; 2000US-0230437P.
PR   06-SEP-2000; 2000US-0230438P.
PR   08-SEP-2000; 2000US-0231242P.
PR   08-SEP-2000; 2000US-0231243P.
PR   08-SEP-2000; 2000US-0231244P.
PR   08-SEP-2000; 2000US-0231413P.
PR   08-SEP-2000; 2000US-0231414P.
PR   08-SEP-2000; 2000US-0232080P.
PR   08-SEP-2000; 2000US-0232081P.
PR   12-SEP-2000; 2000US-0231968P.
PR   14-SEP-2000; 2000US-0232397P.
PR   14-SEP-2000; 2000US-0232398P.
PR   14-SEP-2000; 2000US-0232399P.
PR   14-SEP-2000; 2000US-0232400P.
PR   14-SEP-2000; 2000US-0232401P.
PR   14-SEP-2000; 2000US-0233063P.
PR   14-SEP-2000; 2000US-0233064P.
PR   14-SEP-2000; 2000US-0233065P.
PR   21-SEP-2000; 2000US-0234223P.
PR   21-SEP-2000; 2000US-0234274P.
PR   25-SEP-2000; 2000US-0234997P.
PR   25-SEP-2000; 2000US-0234998P.
PR   26-SEP-2000; 2000US-0235484P.
PR   27-SEP-2000; 2000US-0235834P.
PR   27-SEP-2000; 2000US-0235836P.
PR   29-SEP-2000; 2000US-0236327P.
PR   29-SEP-2000; 2000US-0236367P.
PR   29-SEP-2000; 2000US-0236368P.
PR   29-SEP-2000; 2000US-0236369P.
PR   29-SEP-2000; 2000US-0236370P.
PR   02-OCT-2000; 2000US-0236802P.
PR   02-OCT-2000; 2000US-0237037P.
PR   02-OCT-2000; 2000US-0237038P.
PR   02-OCT-2000; 2000US-0237039P.
PR   02-OCT-2000; 2000US-0237040P.
PR   13-OCT-2000; 2000US-0239935P.
PR   13-OCT-2000; 2000US-0239937P.
PR   20-OCT-2000; 2000US-0240960P.
PR   20-OCT-2000; 2000US-0241221P.
PR   20-OCT-2000; 2000US-0241785P.
PR   20-OCT-2000; 2000US-0241786P.
PR   20-OCT-2000; 2000US-0241787P.
PR   20-OCT-2000; 2000US-0241808P.
PR   20-OCT-2000; 2000US-0241809P.
PR   20-OCT-2000; 2000US-0241826P.
PR   01-NOV-2000; 2000US-0244617P.
PR   08-NOV-2000; 2000US-0246474P.
PR   08-NOV-2000; 2000US-0246475P.
PR   08-NOV-2000; 2000US-0246476P.
PR   08-NOV-2000; 2000US-0246477P.
PR   08-NOV-2000; 2000US-0246478P.
PR   08-NOV-2000; 2000US-0246523P.
PR   08-NOV-2000; 2000US-0246524P.
PR   08-NOV-2000; 2000US-0246525P.
PR   08-NOV-2000; 2000US-0246526P.
PR   08-NOV-2000; 2000US-0246527P.
PR   08-NOV-2000; 2000US-0246528P.
PR   08-NOV-2000; 2000US-0246532P.
PR   08-NOV-2000; 2000US-0246609P.
PR   08-NOV-2000; 2000US-0246610P.
PR   08-NOV-2000; 2000US-0246611P.
PR   08-NOV-2000; 2000US-0246613P.
PR   17-NOV-2000; 2000US-0249207P.
PR   17-NOV-2000; 2000US-0249208P.
PR   17-NOV-2000; 2000US-0249209P.
PR   17-NOV-2000; 2000US-0249210P.
```

```
PR   17-NOV-2000; 2000US-0249211P.
PR   17-NOV-2000; 2000US-0249212P.
PR   17-NOV-2000; 2000US-0249213P.
PR   17-NOV-2000; 2000US-0249214P.
PR   17-NOV-2000; 2000US-0249215P.
PR   17-NOV-2000; 2000US-0249216P.
PR   17-NOV-2000; 2000US-0249217P.
PR   17-NOV-2000; 2000US-0249218P.
PR   17-NOV-2000; 2000US-0249244P.
PR   17-NOV-2000; 2000US-0249245P.
PR   17-NOV-2000; 2000US-0249264P.
PR   17-NOV-2000; 2000US-0249265P.
PR   17-NOV-2000; 2000US-0249297P.
PR   17-NOV-2000; 2000US-0249299P.
PR   17-NOV-2000; 2000US-0249300P.
PR   01-DEC-2000; 2000US-0250160P.
PR   01-DEC-2000; 2000US-0250391P.
PR   05-DEC-2000; 2000US-0251030P.
PR   05-DEC-2000; 2000US-0251988P.
PR   05-DEC-2000; 2000US-0256719P.
PR   06-DEC-2000; 2000US-0251479P.
PR   08-DEC-2000; 2000US-0251856P.
PR   08-DEC-2000; 2000US-0251868P.
PR   08-DEC-2000; 2000US-0251869P.
PR   08-DEC-2000; 2000US-0251989P.
PR   08-DEC-2000; 2000US-0251990P.
PR   11-DEC-2000; 2000US-0254097P.
PR   05-JAN-2001; 2001US-0259678P.
PR   17-JAN-2001; 2001US-00764878.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA, Ruben SM, Barash SC;
XX
DR   WPI; 2003-695900/66.
XX
PT   Novel isolated lung antigen polypeptides useful for treating, preventing,
PT   diagnosing acute myelogenous leukemias, adenocarcinoma, thrombocytopenia,
PT   Von Willebrand's disease.
XX
PS   Disclosure; SEQ ID NO 379; 178pp; English.
XX
CC   The invention relates to an isolated lung antigen polypeptide sequence or
CC   encoded sequence in a cDNA clone. The polypeptide and its polynucleotide
CC   are useful for treating, preventing, diagnosing and/or prognosing
CC   diseases and/or disorders such as pathological cell proliferative
CC   neoplasias e.g. acute myelogenous leukaemias, adenocarcinoma; respiratory
CC   disorders such as chronic rhinitis, sinusitis; immunodeficiencies such as
CC   X-linked agammaglobulinaemia, X-linked infantile agammaglobulinaemia;
CC   inflammatory disorders such as adrenalitis, alveolitis; immune complex
CC   diseases such as serum sickness, polyarteritis nodosa; bleeding disorders
CC   such as thrombocytopenia, Von Willebrand's disease; acquired platelet
CC   dysfunction such as kidney failure, multiple myeloma; disorders
CC   associated with macrophage numbers and/or macrophage function such as
CC   Gaucher's disease, Neimann-Pick disease; tumours such as colon cancer,
CC   pancreatic cancer; renal disorders such as kidney failure, nephritis;
CC   bone disorders such as Albers-Schonberg disease, bowlegs; muscle
CC   disorders such as Becker's muscular dystrophy, Duchenne's muscular
CC   dystrophy; nervous disorders such as ischaemic lesions, traumatic lesions
CC   ; endocrine disorders such as Cushing's syndrome, corticosteroid

  Query Match             4.1%;  Score 19;  DB 10;  Length 32189;
  Best Local Similarity   100.0%;  Pred. No. 26;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         62 CTGATGGTTCAGGAGGCAG 80
              ||||||||||||||||||
Db      23959 CTGATGGTTCAGGAGGCAG 23977


RESULT 6
AAS30113
ID   AAS30113 standard; DNA; 32221 BP.
XX
AC   AAS30113;
XX
DT   21-NOV-2001  (first entry)
XX
DE   Human lung antigen genomic DNA #183.
XX
KW   Lung antigen protein; human; mouse; rabbit; goat; horse; cat; dog;
KW   chicken; sheep; immunosuppressive; antiarthritic; vasotropic;
KW   antirheumatic; antiproliferative; cytostatic; cardiant; neuroprotective;
KW   cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;
KW   ophthalmological; vulnerary; gene therapy; autoimmune disease; neoplasm;
KW   hyperproliferative disorder; breast; liver; cardiovascular disorder; ds;
KW   cerebrovascular disorder; nervous system disorder; bacterial infection;
KW   fungal infection; viral infection; ocular disorder; endocrine disorder;
KW   gastrointestinal disorder; renal disorder; respiratory disorder;
KW   wound healing; skin aging; organ transplantation; food preservative;
KW   tissue regeneration; anti-infertility; food additive.
XX
OS   Homo sapiens.
XX
PN   WO200155303-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US001301.
XX
PR   31-JAN-2000; 2000US-0179065P.
PR   04-FEB-2000; 2000US-0180628P.
PR   24-FEB-2000; 2000US-0184664P.
PR   02-MAR-2000; 2000US-0186350P.
PR   16-MAR-2000; 2000US-0189874P.
PR   17-MAR-2000; 2000US-0190076P.
PR   18-APR-2000; 2000US-0198123P.
PR   19-MAY-2000; 2000US-0205515P.
PR   07-JUN-2000; 2000US-0209467P.
PR   28-JUN-2000; 2000US-0214886P.
PR   30-JUN-2000; 2000US-0215135P.
PR   07-JUL-2000; 2000US-0216647P.
PR   07-JUL-2000; 2000US-0216880P.
PR   11-JUL-2000; 2000US-0217487P.
PR   11-JUL-2000; 2000US-0217496P.
PR   14-JUL-2000; 2000US-0218290P.
PR   26-JUL-2000; 2000US-0220963P.
PR   26-JUL-2000; 2000US-0220964P.
PR   14-AUG-2000; 2000US-0224518P.
PR   14-AUG-2000; 2000US-0224519P.
PR   14-AUG-2000; 2000US-0225213P.
PR   14-AUG-2000; 2000US-0225214P.
PR   14-AUG-2000; 2000US-0225266P.
PR   14-AUG-2000; 2000US-0225267P.
PR   14-AUG-2000; 2000US-0225268P.
PR   14-AUG-2000; 2000US-0225270P.
PR   14-AUG-2000; 2000US-0225447P.
PR   14-AUG-2000; 2000US-0225757P.
PR   14-AUG-2000; 2000US-0225758P.
PR   14-AUG-2000; 2000US-0225759P.
PR   18-AUG-2000; 2000US-0226279P.
PR   22-AUG-2000; 2000US-0226681P.
PR   22-AUG-2000; 2000US-0226868P.
PR   22-AUG-2000; 2000US-0227182P.
PR   23-AUG-2000; 2000US-0227009P.
PR   30-AUG-2000; 2000US-0228924P.
PR   01-SEP-2000; 2000US-0229287P.
PR   01-SEP-2000; 2000US-0229343P.
PR   01-SEP-2000; 2000US-0229344P.
PR   01-SEP-2000; 2000US-0229345P.
PR   05-SEP-2000; 2000US-0229509P.
PR   05-SEP-2000; 2000US-0229513P.
PR   06-SEP-2000; 2000US-0230437P.
PR   06-SEP-2000; 2000US-0230438P.
PR   08-SEP-2000; 2000US-0231242P.
PR   08-SEP-2000; 2000US-0231243P.
```

PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-457723/49.
XX
PT Isolated polypeptide for treating, preventing and/ or prognosing
PT respiratory disorders related to the lung including lung cancers and also
PT for testing and detection e.g. diagnosis.
XX
PS Claim 1; SEQ ID NO 377; 507pp; English.
XX
CC Sequences AAS29931-AAS30164 represent genomic DNA molecules, which encode
CC the lung antigen polypeptides of the invention. Lung antigen polypeptides
CC and their associated polynucleotides are useful in the diagnosis,
CC treatment and prevention of various types of disorders in e.g. humans,
CC mice, rabbits, goats, horses, cats, dogs, chickens or sheep. A
CC pathological condition can be determined by detecting the presence or
CC absence of a mutation in a lung antigen polynucleotide. The treatable
CC disorders include autoimmune diseases such as rheumatoid arthritis,
CC hyperproliferative disorders such as neoplasms of the breast or liver,
CC cardiovascular disorders such as cardiac arrest, cerebrovascular
CC disorders such as cerebral ischaemia, nervous system disorders such as
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi,
CC ocular disorders such as corneal infection, endocrine disorders such as
CC premature labour and infertility, gastrointestinal disorders such as
CC Crohn's disease, renal disorders such as glomerulonephritis and
CC respiratory disorders such as asthma and pleurisy. The polypeptides can
CC also be used to aid wound healing, to prevent skin aging due to sunburn,
CC to maintain organs before transplantation, to regenerate tissues and in
CC chemotaxis. The polypeptides can also be used as a food additive or
CC preservative to increase or decrease storage capabilities. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences

Query Match 4.1%; Score 19; DB 5; Length 32221;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 62 CTGATGGTTCAGGAGGCAG 80
||||||||||||||||||||
Db 23989 CTGATGGTTCAGGAGGCAG 24007

RESULT 7
ADB33450
ID ADB33450 standard; DNA; 32221 BP.
XX
AC ADB33450;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human novel lung related polypeptide DNA SEQ ID NO 377.
XX

KW gene therapy; lung antigen; neoplasia; acute myelogenous leukaemia;
KW adenocarcinoma; respiratory disorder; chronic rhinitis; sinusitis;
KW immunodeficiency; X-linked agammaglobulinaemia;
KW X-linked infantile agammaglobulinaemia; inflammatory disorder;
KW adrenalitis; alveolitis; immune complex disease; serum sickness;
KW polyarteritis nodosa; bleeding disorder; thrombocytopenia;
KW Von Willebrand's disease; acquired platelet dysfunction; kidney failure;
KW multiple myeloma; macrophage related disorder; Gaucher's disease;
KW Neimann-Pick disease; tumour; colon cancer; pancreatic cancer;
KW renal disorder; nephritis; bone disorder; Albers-Schonberg disease;
KW bowleg; muscle disorder; Becker's muscular dystrophy;
KW Duchenne's muscular dystrophy; nervous disorder; ischaemic lesion;
KW traumatic lesion; endocrine disorder; Cushing's syndrome;
KW corticosteroid deficiency; gastrointestinal disorder; dysphagia;
KW gastric reflux; human; ds.
XX
OS Homo sapiens.
XX
PN US2003054368-A1.
XX
PD 20-MAR-2003.
XX
PF 22-FEB-2002; 2002US-00079854.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.

PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764878.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM, Barash SC;
XX
DR WPI; 2003-695900/66.
XX
PT Novel isolated lung antigen polypeptides useful for treating, preventing,
PT diagnosing acute myelogenous leukemias, adenocarcinoma, thrombocytopenia,
PT Von Willebrand's disease.
XX
PS Disclosure; SEQ ID NO 377; 178pp; English.
XX
CC The invention relates to an isolated lung antigen polypeptide sequence or
CC encoded sequence in a cDNA clone. The polypeptide and its polynucleotide
CC are useful for treating, preventing, diagnosing and/or prognosing
CC diseases and/or disorders such as pathological cell proliferative
CC neoplasias e.g. acute myelogenous leukaemias, adenocarcinoma; respiratory
CC disorders such as chronic rhinitis, sinusitis; immunodeficiencies such as
CC X-linked agammaglobulinaemia, X-linked infantile agammaglobulinaemia;
CC inflammatory disorders such as adrenalitis, alveolitis; immune complex
CC diseases such as serum sickness, polyarteritis nodosa; bleeding disorders
CC such as thrombocytopenia, Von Willebrand's disease; acquired platelet
CC dysfunction such as kidney failure, multiple myeloma; disorders
CC associated with macrophage numbers and/or macrophage function such as
CC Gaucher's disease, Neimann-Pick disease; tumours such as colon cancer,
CC pancreatic cancer; renal disorders such as kidney failure, nephritis;
CC bone disorders such as Albers-Schonberg disease, bowlegs; muscle
CC disorders such as Becker's muscular dystrophy, Duchenne's muscular
CC dystrophy; nervous disorders such as ischaemic lesions, traumatic lesions
CC ; endocrine disorders such as Cushing's syndrome, corticosteroid

Query Match 4.1%; Score 19; DB 10; Length 32221;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 62 CTGATGGTTCAGGAGGCAG 80
||||||||||||||||||||
Db 23989 CTGATGGTTCAGGAGGCAG 24007

RESULT 8
ABN77046/c
ID ABN77046 standard; cDNA; 278 BP.
XX
AC ABN77046;
XX
DT 08-JUL-2002 (first entry)
XX
DE Human ORF1993 cDNA, SEQ ID NO:3985.
XX
KW Human; ORF; open reading frame; ORFX; drug screening; diagnosis;
KW disease monitoring; cytokine; cell proliferation; cell differentiation;
KW immune modulation; haematopoiesis regulation; tissue growth;
KW angiogenesis; activin; inhibin; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; tumour inhibition; bodily characteristic; fertility;
KW behaviour; cancer; proliferative disorder; neurological disorder;
KW cardiovascular disease; immune system disorder; organ transplantation;
KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
KW hypothyroidism; cholesterol ester storage disease; infection; vulnerary;
KW vasotropic; antipsoriatic; antidiabetic; cytostatic; nootropic;
KW neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;
KW cardiant; hypotensive; antithyroid; antiinflammatory; immunomodulator;
KW dermatological; analgesic; virucide; antibacterial; fungicide; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200190366-A2.
XX
PD 29-NOV-2001.
XX
PF 24-MAY-2001; 2001WO-US017076.
XX
PR 24-MAY-2000; 2000US-0206690P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Leach MD, Shimkets RA;
XX
DR WPI; 2002-106200/14.
DR P-PSDB; ABP33020.
XX
PT Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and disorders related to organ
PT transplantation.
XX
PS Claim 1; Page 1253; 2508pp; English.
XX
CC Sequences ABP31028-ABP35561 represent 4534 novel human proteins
CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-
CC ABN79587 represent cDNAs encoding them. The invention also encompasses
CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
CC referred to as ORFX) proteins, polynucleotides at least 85% identical to
CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
CC polynucleotides, the recombinant production of ORFX proteins, antibodies
CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and
CC polypeptides, methods of screening for modulators of ORFX expression or
CC activity, and methods of screening individuals for a predisposition to an
CC ORFX-associated disorder. The ORFX proteins of the invention have a wide
CC range of biological activities, such as cytokine, cell proliferation,
CC cell differentiation, immune modulation, haematopoiesis regulation,
CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
CC chemokinetic activity, haemostatic activity, thrombolytic activity,
CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,
CC and antiinfective activity, and may also be involved in the determination
CC of bodily characteristics, fertility and behaviour. ORFX proteins,
CC nucleic acids and antibodies may be used in the treatment of cancers,
CC other proliferative disorders such as psoriasis and benign tumours,
CC neurological disorders such as epilepsy and Alzheimer's disease,
CC cardiovascular diseases, immune system disorders, disorders related to
CC organ transplantation, disorders of tissue growth and regeneration,
CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
CC storage disease, and infectious diseases caused by viral, bacterial,
CC fungal and other pathogens. ORFX nucleic acids may also be used as a
CC source of primers and probes, in the detection of ORFX genomic sequences
CC or transcripts, in the identification and cloning of homologous
CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
CC nucleic acids may additionally be used to produce transgenic animals
CC which may be useful for studying the function and/or activity of ORFX
CC protein, and in drug screening. The ORFX proteins may also be used as
CC immunogens to generate specific antibodies, which are useful in the
CC diagnosis, treatment and monitoring of ORFX-associated diseases
XX
SQ Sequence 278 BP; 73 A; 74 C; 79 G; 52 T; 0 U; 0 Other;

Query Match 3.8%; Score 18; DB 6; Length 278;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 83 GCAGTACCAGCTTCTGCA 100

```
                  ||||||||||||||||||
Db        126 GCAGTACCAGCTTCTGCA 109


RESULT 9
ADL24711
ID   ADL24711 standard; DNA; 309 BP.
XX
AC   ADL24711;
XX
DT   20-MAY-2004  (first entry)
XX
DE   Intestinal epithelium/peyer's patch M cell-associated DNA sequence #72.
XX
KW   intestinal epithelium cell development; peyer's patch M cell development;
KW   inflammatory bowel disease; glutenenteropathy; infectious disease;
KW   autoimmune disease; haemolytic anaemia; rheumatoid arthritis; dermatitis;
KW   Grave's disease; multiple sclerosis; allergy; asthma; diabetic mellitus;
KW   immune system disorder; hypersensitivity; anaphylaxis;
KW   blood group incompatibility; ds; mouse; murine.
XX
OS   Mus musculus.
XX
PN   WO200280852-A2.
XX
PD   17-OCT-2002.
XX
PF   04-APR-2002; 2002WO-US010873.
XX
PR   04-APR-2001; 2001US-0281416P.
XX
PA   (DIGI-) DIGITAL GENE TECHNOLOGIES INC.
XX
PI   Brayden DJ,  Byrne D,  O'mahony DJ,  Evans CF,  Mah SP,  Lo DD;
XX
DR   WPI; 2003-075470/07.
XX
PT   Novel isolated or purified polypeptide encoded by genes associated with
PT   intestinal epithelium or M cell development, differentiation or function,
PT   useful for treating autoimmune diseases and infectious diseases.
XX
PS   Claim 1; SEQ ID NO 221; 152pp; English.
XX
CC   The invention comprises DNA sequences which are associated with
CC   intestinal epithelium and peyer's patch M cells. The DNA sequences of the
CC   invention are useful for assessing, modifying, modulating or regulating
CC   intestinal epithelium or M cell development. The DNA sequences of the
CC   invention are also useful in the treatment of: inflammatory bowel
CC   disease, glutenenteropathy, infectious diseases, autoimmune diseases
CC   (e.g. haemolytic anaemia, rheumatoid arthritis, dermatitis, Grave's
CC   disease, multiple sclerosis, allergy, asthma and diabetic mellitus),
CC   diseases or disorders of the immune system, hypersensitivity,
CC   anaphylaxis, and blood group incompatibility. The present nucleic acid
CC   represents an intestinal epithelium/peyer's patch M cell-associated DNA
CC   sequence of the invention. NOTE: The present sequence is not shown in the
CC   specification, but has been retrieved from the WIPO website.
XX
SQ   Sequence 309 BP; 98 A; 62 C; 69 G; 80 T; 0 U; 0 Other;

  Query Match             3.8%;  Score 18;  DB 10;  Length 309;
  Best Local Similarity   100.0%;  Pred. No. 95;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         69 TTCAGGAGGCAGGGGCAG 86
              ||||||||||||||||||
Db         43 TTCAGGAGGCAGGGGCAG 60


RESULT 10
ACF85026
ID   ACF85026 standard; DNA; 459 BP.
XX
AC   ACF85026;
XX
DT   02-JUN-2005  (first entry)
XX
DE   Human SIRS/sepsis diagnostic marker DNA fragment 3886.
XX
KW   Systemic inflammatory response syndrome; SIRS; antibacterial;
KW   immunosuppressive; antiinflammatory; diagnosis; sepsis; ds.
XX
OS   Homo sapiens.
XX
PN   WO2004087949-A2.
XX
PD   14-OCT-2004.
XX
PF   31-MAR-2004; 2004WO-EP003419.
XX
PR   02-APR-2003; 2003DE-01015031.
PR   08-AUG-2003; 2003DE-01036511.
PR   02-SEP-2003; 2003DE-01040395.
XX
PA   (SIRS-) SIRS LAB GMBH.
XX
PI   Russwurm S,  Reinhart K,  Saluz H,  Straube E,  Zipfel PF,  Deigner H;
XX
DR   WPI; 2004-748070/73.
XX
PT   In vitro detection of systemic inflammatory response syndrome and related
PT   conditions, for e.g. monitoring progression, comprises detecting abnormal
PT   expression of disease-related genes.
XX
PS   Disclosure; Page; 75pp; German.
XX
CC   The invention relates to a novel method for in vitro detection of
CC   systemic inflammatory response syndrome (SIRS). The method comprises
CC   detecting abnormal expression of disease-related genes, or their
CC   associated peptides. The method of the invention demonstrates
CC   antibacterial, immunosuppressive and antiinflammatory applications and
CC   may be used for early differential diagnosis, monitoring progression,
CC   assessing risk, assessing the likely response to treatment and for post
CC   mortem diagnosis of systemic inflammatory response syndrome, sepsis and
CC   sepsis-like conditions. The recombinant or synthetic nucleic acid
CC   sequences of the invention, or derived proteins or peptides, may be
CC   useful as calibrants in assays for the specified diseases, for evaluating
CC   activity or toxicity in screening for active agents and/or for
CC   preparation of agents for treatment or prevention of the specified
CC   diseases. The current sequence is that of a human SIRS/sepsis diagnostic
CC   marker DNA fragment of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from WIPO at ftp.wipo.int/pub/published
CC   pct_sequences. Furthermore, a number of arbitrary SEQ ID NO.s are
CC   disclosed within the specification, however, these have not been taken
CC   into account during indexing due to inconsistencies in application and
CC   format
XX
SQ   Sequence 459 BP; 145 A; 82 C; 87 G; 145 T; 0 U; 0 Other;

  Query Match             3.8%;  Score 18;  DB 13;  Length 459;
  Best Local Similarity   100.0%;  Pred. No. 94;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        327 GTAATGATGTTTTCTTCT 344
              ||||||||||||||||||
Db        244 GTAATGATGTTTTCTTCT 261


RESULT 11
ABV46132/c
ID   ABV46132 standard; cDNA; 465 BP.
XX
AC   ABV46132;
XX
DT   16-SEP-2002  (first entry)
XX
```

DE Human prostate expression marker cDNA 46123.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodyanamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 9110-9111; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodyanamic or pharmacogenomic marker
XX
SQ Sequence 465 BP; 161 A; 79 C; 82 G; 143 T; 0 U; 0 Other;

Query Match 3.8%; Score 18; DB 5; Length 465;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 185 GAGTTGTTTATATATTTC 202
||||||||||||||||||
Db 428 GAGTTGTTTATATATTTC 411

RESULT 12
AAC35341/c
ID AAC35341 standard; DNA; 565 BP.
XX
AC AAC35341;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 9835.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.

PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 3.8%; Score 18; DB 3; Length 565;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        397 TGTTTTTGAGTTCTTCAA 414
              ||||||||||||||||||
Db         57 TGTTTTTGAGTTCTTCAA 40
```

RESULT 13
AEF29701/c
ID AEF29701 standard; cDNA; 565 BP.
XX
AC AEF29701;
XX
DT 23-MAR-2006 (first entry)
XX
DE Lead_CeresClone cDNA SEQ ID NO:1901.
XX
KW ss; gene; plant; transgenic plant; crop improvement;
KW abiotic stress tolerance; plant growth regulation.
XX
OS Arabidopsis thaliana.
XX
PN WO2006004955-A2.
XX
PD 12-JAN-2006.
XX
PF 30-JUN-2005; 2005WO-US023326.
XX
PR 30-JUN-2004; 2004US-0583621P.
PR 30-JUN-2004; 2004US-0584800P.
PR 30-JUN-2004; 2004US-0584829P.

XX
PA (CERE-) CERES INC.
XX
PI Alexandrov N, Brover V, Mascia P, Feldmann K;
XX
DR WPI; 2006-090599/09.
DR P-PSDB; AEF29702.
XX
PT New isolated nucleic acid molecule modifying plant phenotypes and
PT characteristics and the polypeptide it encodes, useful for making
PT transgenic plants with improved characteristics.
XX
PS Claim 2; SEQ ID NO 1901; 612pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule modifying
CC plant phenotypes and characteristics, comprising a nucleotide sequence
CC that encodes an amino acid sequence exhibiting at least 85% sequence
CC identity to an amino acid sequence in the sequence listing or in the
CC ortholog alignments of Figure 1, a nucleic acid, which is a complement of
CC (a), a nucleic acid, which is the reverse of the nucleotide sequence in
CC (a)(such that the reverse nucleotide sequence has a sequence order which
CC is the reverse of the sequence order of (a)) or a nucleic acid capable of
CC hybridizing (a-c), under conditions that permit formation of a nucleic
CC acid duplex at a temperature of 40-48 degrees C below the melting
CC temperature of the nucleic acid duplex. Also included are a vector
CC construct (comprising a first nucleic acid having a regulatory sequence
CC capable of causing transcription and/or translation in a plant, operably
CC linked to a second nucleic acid having the sequence of the isolated
CC nucleic acid molecule), a host cell comprising the isolated nucleic acid
CC molecule that is flanked by exogenous sequence, a host cell comprising
CC the vector construct, an isolated polypeptide comprising an amino acid
CC sequence exhibiting at least 85% sequence identity to those cited above,
CC introducing an isolated nucleic acid into a host cell, transforming a
CC host cell, detecting a nucleic acid in a sample, a host cell or organism
CC comprising the nucleic acid molecule, a plant generated from the plant
CC cell or seed, a plant(plant cell, plant material or seed) comprising the
CC nucleic acid molecule (where the plant has improved characteristics as
CC compared to a wild type plant), improving plant characteristics in a
CC plant comprising transforming the plant with the nucleic acid sequence,
CC and a transgenic plant having a gene construct (comprising the nucleic
CC encoding a component operably linked to a plant promoter so that the
CC component is ectopically overexpressed in the transgenic plant). The
CC transgenic plant exhibits faster rate of growth, greater fresh of dry
CC weight of maturation, greater fruit or seed yield, higher tolerance to
CC pH, higher tolerance to low phosphate concentration, or higher tolerance
CC to low nitrogen concentration than a progenitor plant, which does not
CC contain the progenitor construct, when the transgenic plant and
CC progenitor plant are cultivated under identical environmental conditions,
CC where the component is any one of the polypeptides cited above. The
CC nucleic acid molecules are useful for producing transgenic plants with
CC improved characteristics. The present sequence is a plant nucleic acid
CC (cDNA) of the invention.
XX
SQ Sequence 565 BP; 210 A; 99 C; 108 G; 148 T; 0 U; 0 Other;

Query Match 3.8%; Score 18; DB 15; Length 565;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 397 TGTTTTTGAGTTCTTCAA 414
||||||||||||||||||
Db 57 TGTTTTTGAGTTCTTCAA 40

RESULT 14
AAH12181/c
ID AAH12181 standard; cDNA; 575 BP.
XX
AC AAH12181;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA clone (3'-primer) SEQ ID NO:9016.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
PS Claim 3; SEQ ID NO 9016; 2537pp + Sequence Listing; English.
XX
CC The present invention describes primer sets for synthesising 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
SQ Sequence 575 BP; 184 A; 104 C; 124 G; 155 T; 0 U; 8 Other;

Query Match 3.8%; Score 18; DB 4; Length 575;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 331 TGATGTTTTCTTCTCAAC 348
||||||||||||||||||
Db 90 TGATGTTTTCTTCTCAAC 73

RESULT 15
AAA70256
ID AAA70256 standard; DNA; 675 BP.
XX
AC AAA70256;
XX
DT 07-NOV-2000 (first entry)
XX

DE Plasmodium falciparum chromosome 2 related DNA sequence SEQ ID NO:389.
XX
KW Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;
KW antimalarial; malaria; protozoacide; infection; insecticide; ds.
XX
OS Plasmodium falciparum.
XX
PN WO200025728-A2.
XX
PD 11-MAY-2000.
XX
PF 05-NOV-1999; 99WO-US026796.
XX
PR 05-NOV-1998; 98US-0107131P.
XX
PA (HOFF/) HOFFMAN S.
PA (CARU/) CARUCCI D.
PA (GARD/) GARDNER M.
PA (VENT/) VENTER J C.
XX
PI Hoffman S, Carucci D, Gardner M, Venter JC;
XX
DR WPI; 2000-365347/31.
XX
PT Proteins encoded by chromosome 2 of the human malarial parasite,
PT Plasmodium falciparum, useful as antimalarial vaccines and in the
PT diagnosis of P.falciparum infection.
XX
PS Disclosure; Page 564; 577pp; English.
XX
CC The present invention describes proteins and their fragments (I) encoded
CC by chromosome 2 of the human malarial parasite, Plasmodium falciparum.
CC Also described are: (1) nucleotide sequences (II) encoding (I); and (2)
CC vaccines against P. falciparum infection comprising (I) or (II). (I) and
CC (II) are useful for the development of vaccines against P. falciparum
CC infection. (I) and polyclonal antisera or a monoclonal antibody raised to
CC immunogens comprising the sequences of (I), are useful in the detection
CC of infection with P. falciparum. Furthermore, (I) (especially when they
CC are rifins or secreted or membrane proteins) can aid the identification
CC of drugs to treat or prevent P. falciparum infection, or they can be used
CC to identify drug resistance in P. falciparum. Sequencing of the
CC Plasmodium chromosome 2 and the subsequent identification of proteins
CC encoded by it will help to expand our understanding of parasite biology,
CC a process hampered by the complexity of the parasitic lifecycle, and
CC provide new targets for vaccine and drug development. Parasite resistance
CC to drugs and mosquito resistance to insecticides have led to a resurgence
CC of malaria in many parts of the world, and there is a pressing need for
CC vaccines and new drugs. AAA70078 to AAA70287 and AAB18144 to AAB18352
CC represent nucleotide and protein sequences given in the present
CC invention, but which are not specifically mentioned within the
CC specification
XX
SQ Sequence 675 BP; 300 A; 71 C; 91 G; 213 T; 0 U; 0 Other;

Query Match 3.8%; Score 18; DB 3; Length 675;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GAAAATAAATGGAACACA 19
||||||||||||||||||
Db 625 GAAAATAAATGGAACACA 642

Search completed: November 3, 2006, 20:46:05
Job time : 279.593 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 3, 2006, 20:46:30 ; Search time 2594.63 Seconds
(without alignments)
11534.376 Million cell updates/sec

Title: US-10-764-316-5
Perfect score: 468
Sequence: 1 tgaaaataaatggaacacac.....tggatgattgcaggagattt 468

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 6366136 seqs, 31973710525 residues

Word size : 1

Total number of hits satisfying chosen parameters: 12730834

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl:*
1: gb_env:*
2: gb_pat:*
3: gb_ph:*
4: gb_pl:*
5: gb_pr:*
6: gb_ro:*
7: gb_sts:*
8: gb_sy:*
9: gb_un:*
10: gb_vi:*
11: gb_ov:*
12: gb_htg:*
13: gb_in:*
14: gb_om:*
15: gb_ba:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	467	99.8	1991	15	AF373206	AF373206 Escherich
	2	467	99.8	2305	15	ECCDTABC	U03293 Escherichia
	3	462	98.7	462	15	AY351905	AY351905 Escherich
	4	462	98.7	462	15	AY351906	AY351906 Escherich
	5	462	98.7	462	15	AY351907	AY351907 Escherich
	6	302	64.5	494	15	AY426341	AY426341 Escherich
	7	298	63.7	463	15	AF373205	AF373205 Escherich
	8	185	39.5	495	15	AY426340	AY426340 Escherich
	9	44	9.4	822	2	AX140193	AX140193 Sequence
	10	44	9.4	10241	15	AY578329	AY578329 Escherich
	11	41	8.8	417	15	AY162217	AY162217 Escherich
c	12	21	4.5	136195	11	BX465849	BX465849 Zebrafish
	13	21	4.5	141587	12	AC128885	AC128885 Rattus no
c	14	21	4.5	191669	12	AC175073	AC175073 Bos tauru
c	15	21	4.5	238377	12	AC107439	AC107439 Rattus no
	16	20	4.3	36255	12	AC016208	AC016208 Homo sapi
c	17	20	4.3	59281	5	AC084872	AC084872 Homo sapi
c	18	20	4.3	83102	2	CQ869791	CQ869791 Sequence
c	19	20	4.3	110000	4	CR382132_15	Continuation (16 o
c	20	20	4.3	136072	12	CT030674	CT030674 Platynere
c	21	20	4.3	146770	6	AC113992	AC113992 Mus muscu
c	22	20	4.3	154679	5	AC147086	AC147086 Pan trogl
	23	20	4.3	157886	5	AC146051	AC146051 Pan trogl
c	24	20	4.3	166703	5	AL139378	AL139378 Human DNA
c	25	20	4.3	179223	5	AC108709	AC108709 Homo sapi
	26	20	4.3	179676	12	AC012221	AC012221 Homo sapi
c	27	20	4.3	206915	6	AC110248	AC110248 Mus muscu
	28	20	4.3	207977	12	AC161965	AC161965 Bos tauru
c	29	20	4.3	216816	12	CR759846	CR759846 Danio rer
	30	20	4.3	233498	12	AC015882	AC015882 Homo sapi
	31	20	4.3	242358	12	AC165857	AC165857 Bos tauru
c	32	20	4.3	243957	12	AC096118	AC096118 Rattus no
c	33	20	4.3	244813	12	AC153211	AC153211 Bos tauru
	34	20	4.3	257317	12	AC157305	AC157305 Bos tauru
c	35	20	4.3	258016	12	AC163323	AC163323 Bos tauru
c	36	20	4.3	271832	12	AC156493	AC156493 Bos tauru
c	37	20	4.3	304610	12	AC169320	AC169320 Bos tauru
c	38	19	4.1	776	7	BV613747	BV613747 S215P6074
	39	19	4.1	934	11	SAU418605	AJ418605 Sparus au
c	40	19	4.1	16922	14	AJ639873	AJ639873 Potorous
	41	19	4.1	44069	11	BX649362	BX649362 Zebrafish
	42	19	4.1	63441	12	AC173964	AC173964 Medicago
c	43	19	4.1	70385	12	AC101370	AC101370 Mus muscu
	44	19	4.1	72820	5	AC135849	AC135849 Pan trogl
	45	19	4.1	76175	4	T1N6	AC009273 Sequence

ALIGNMENTS

RESULT 1
AF373206
LOCUS AF373206 1991 bp DNA linear BCT 22-JUL-2005
DEFINITION Escherichia coli strain MBU. E 412 cytolethal distending toxin B (cdtB) gene, complete cds; and unknown genes.
ACCESSION AF373206
VERSION AF373206.4 GI:49912092
KEYWORDS .
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.
REFERENCE 1 (bases 1 to 1991)
AUTHORS Bouzari,S., Oloomi,M. and Oswald,E.
TITLE Detection of the cytolethal distending toxin locus cdtB among diarrheagenic Escherichia coli isolates from humans in Iran
JOURNAL Res. Microbiol. 156 (2), 137-144 (2005)
PUBMED 15748977
REFERENCE 2 (bases 1 to 1991)
AUTHORS Bouzari,S., Oloomi,M. and Zarepoor,M.
TITLE Direct Submission
JOURNAL Submitted (22-APR-2001) Molecular Biology, Institute Pasteur of Iran, Pasteur Ave., Tehran 13164, Iran
REFERENCE 3 (bases 1 to 1991)
AUTHORS Bouzari,S., Oloomi,M. and Zarepoor,M.
TITLE Direct Submission
JOURNAL Submitted (14-JUN-2004) Molecular Biology, Institute Pasteur of Iran, Pasteur Ave., Tehran 13164, Iran
REMARK Sequence update by submitter
REFERENCE 4 (bases 1 to 1991)
AUTHORS Bouzari,S., Oloomi,M. and Zarepoor,M.
TITLE Direct Submission
JOURNAL Submitted (24-JUN-2004) Molecular Biology, Institute Pasteur of Iran, Pasteur Ave., Tehran 13164, Iran
REMARK Sequence update by submitter
REFERENCE 5 (bases 1 to 1991)
AUTHORS Bouzari,S., Oloomi,M. and Zarepoor,M.
TITLE Direct Submission
JOURNAL Submitted (07-JUL-2004) Molecular Biology, Institute Pasteur of Iran, Pasteur Ave., Tehran 13164, Iran
REMARK Sequence update by submitter

```
COMMENT     On Jul 7, 2004 this sequence version replaced gi:49175456.
FEATURES             Location/Qualifiers
     source          1. .1991
                     /organism="Escherichia coli"
                     /mol_type="genomic DNA"
                     /strain="MBU. E 412"
                     /db_xref="taxon:562"
     CDS             1. .675
                     /note="ORF A"
                     /codon_start=1
                     /transl_table=11
                     /product="unknown"
                     /protein_id="AAT65834.2"
                     /db_xref="GI:49912093"
                     /translation="MTGCSNGIGDSPSPPGKNVELVGIPGQGIAVTSNGATPTLGANN
                     TEFPEVSIMSTGGALLTIWARPVRNWLWGYTPFDSVNFGENRNWKVVDGKDAGTVKFV
                     NVAQGTCMEAFKNGVIHNTCDDNSLSQEFQLLPSTNGNVLIRSSALQTCIRADYLSRT
                     ILSPFAFTITLEKCPGAKEETQEMLWAISPPVRAAKPNLIKPELRPFRPLPIPPHDKP
                     DGMEGV"
     gene            672. .1493
                     /gene="cdtB"
     CDS             672. .1493
                     /gene="cdtB"
                     /note="CdtB"
                     /codon_start=1
                     /transl_table=11
                     /product="cytolethal distending toxin B"
                     /protein_id="AAM21324.3"
                     /db_xref="GI:49175457"
                     /translation="MKKLLFLLMILPGISFADLSDFKVATWNLQGSNAPTENKWNTHV
                     RQLVTGSGAVDILMVQEAGAVPASATLTEREFSTPGIPMNEYIWNTGTNSRPQELFIY
                     FSRVDAFANRVNLAIVSNRRADEVIVLPPPTVVSRPIIGIRIGNDVFFSTHALANRGV
                     DSGAIVNSVFEFFNRQTDPIRQAANWMIAGDFNRSPATLFSTLEPGIRNHVNIIAPPD
                     PTQASGGVLDYAVVGNSVSFVLPLLRASLLFGLLRGQIASDHFPVGFIPGRGARR"
     CDS             1490. .>1991
                     /note="ORF C"
                     /codon_start=1
                     /transl_table=11
                     /product="unknown"
                     /protein_id="AAT68871.1"
                     /db_xref="GI:49912094"
                     /translation="MKTVIVFFVLLLTGCASEPANQRNLLTQFVGDNAPVDPEPSPVL
                     VNIRNVLTGGIIRNPVGSDFNVNNWVISEVKTNDLDLISAPGGHVQIKNPDGNECFAI
                     LNGQLAVAKQCSESDRNALFTFITSDTGAVQIKSIGSGQCLWNGESITDFRLKKCVDD
                     LGRPFDT"

ORIGIN

  Query Match          99.8%;  Score 467;  DB 15;  Length 1991;
  Best Local Similarity  100.0%;  Pred. No. 1.5e-291;
  Matches  467;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        2 GAAAATAAATGGAACACACATGTCCGACAACTTGTTACGGGAAGTGGTGCTGTTGATATC 61
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      780 GAAAATAAATGGAACACACATGTCCGACAACTTGTTACGGGAAGTGGTGCTGTTGATATC 839

Qy       62 CTGATGGTTCAGGAGGCAGGGGCAGTACCAGCTTCTGCAACGTTGACTGAGCGAGAATTT 121
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      840 CTGATGGTTCAGGAGGCAGGGGCAGTACCAGCTTCTGCAACGTTGACTGAGCGAGAATTT 899

Qy      122 AGCACTCCTGGTATTCCGATGAATGAGTATATCTGGAATACCGGAACCAATAGTCGCCCA 181
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      900 AGCACTCCTGGTATTCCGATGAATGAGTATATCTGGAATACCGGAACCAATAGTCGCCCA 959

Qy      182 CAGGAGTTGTTTATATATTTCTCACGTGTTGATGCATTCGCTAACAGAGTAAATCTTGCG 241
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      960 CAGGAGTTGTTTATATATTTCTCACGTGTTGATGCATTCGCTAACAGAGTAAATCTTGCG 1019

Qy      242 ATTGTTTCAAACAGAAGAGCTGATGAGGTGATTGTATTACCTCCTCCAACTGTTGTATCA 301
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1020 ATTGTTTCAAACAGAAGAGCTGATGAGGTGATTGTATTACCTCCTCCAACTGTTGTATCA 1079

Qy      302 CGACCGATCATCGGCATTAGAATTGGTAATGATGTTTTCTTCTCAACCCATGCATTGGCG 361
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db     1080 CGACCGATCATCGGCATTAGAATTGGTAATGATGTTTTCTTCTCAACCCATGCATTGGCG 1139

Qy      362 AATCGGGGCGTGGATTCAGGAGCAATTGTAAATAGTGTTTTTGAGTTCTTCAACAGACAA 421
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1140 AATCGGGGCGTGGATTCAGGAGCAATTGTAAATAGTGTTTTTGAGTTCTTCAACAGACAA 1199

Qy      422 ACGGATCCTATAAGACAGGCCGCTAACTGGATGATTGCAGGAGATTT 468
          |||||||||||||||||||||||||||||||||||||||||||||||
Db     1200 ACGGATCCTATAAGACAGGCCGCTAACTGGATGATTGCAGGAGATTT 1246


RESULT 2
ECCDTABC
LOCUS       ECCDTABC                2305 bp    DNA     linear   BCT 29-JAN-1999
DEFINITION  Escherichia coli E6468/62 (O86:H34) cytolethal distending toxin
            (cdtA, cdtB, cdtC) genes, complete cds.
ACCESSION   U03293
VERSION     U03293.1  GI:416213
KEYWORDS    .
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1  (bases 1 to 2305)
  AUTHORS   Scott,D.A. and Kaper,J.B.
  TITLE     Cloning and sequencing of the genes encoding Escherichia coli
            cytolethal distending toxin
  JOURNAL   Infect. Immun. 62 (1), 244-251 (1994)
   PUBMED   8262635
REFERENCE   2  (bases 1 to 2305)
  AUTHORS   Scott,D.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-NOV-1993) Scott D.A., University of Maryland School
            of Medicine, Geographic Med., Center for Vaccine Dev., 10 South
            Pine Street, Baltimore, MD 21201, USA
FEATURES             Location/Qualifiers
     source          1. .2305
                     /organism="Escherichia coli"
                     /mol_type="genomic DNA"
                     /strain="E6468/62 (O86:H34)"
                     /specific_host="Homo sapiens"
                     /db_xref="taxon:562"
                     /clone="pDS7.96"
     gene            201. .914
                     /gene="cdtA"
     CDS             201. .914
                     /gene="cdtA"
                     /experiment="experimental evidence, no additional details
                     recorded"
                     /note="GTG start codon"
                     /codon_start=1
                     /transl_table=11
                     /product="CdtA"
                     /protein_id="AAD10621.1"
                     /db_xref="GI:416214"
                     /translation="MDKKLIAFLCTLIITGCSNGIGDSPSPPGKNVELVGIPGQGIAV
                     TSNGATPTLGANNTEFPEVSIMSTGGALLTIWARPVRNWLWGYTPFDSVNFGENRNWK
                     VVDGKDAGTVKFVNVAQGTCMEAFKNGVIHNTCDDNSLSQEFQLLPSTNGNVLIRSSA
                     LQTCIRADYLSRTILSPFAFTITLEKCPGAKEETQEMLWAISPPVRAAKPNLIKPELR
                     PFRPLPIPPHDKPDGMEGV"
     gene            911. .1732
                     /gene="cdtB"
     CDS             911. .1732
                     /gene="cdtB"
                     /experiment="experimental evidence, no additional details
                     recorded"
                     /codon_start=1
                     /transl_table=11
                     /product="CdtB"
                     /protein_id="AAD10622.1"
                     /db_xref="GI:416215"
                     /translation="MKKLLFLLMILPGISFADLSDFKVATWNLQGSNAPTENKWNTHV
                     RQLVTGSGAVDILMVQEAGAVPASATLTEREFSTPGIPMNEYIWNTGTNSRPQELFIY
```

```
                     FSRVDAFANRVNLAIVSNRRADEVIVLPPPTVVSRPIIGIRIGNDVFFSTHALANRGV
                     DSGAIVNSVFEFFNRQTDPIRQAANWMIAGDFNRSPATLFSTLEPGIRNHVNIIAPPD
                     PTQASGGVLDYAVVGNSVSFVLPLLRASLLFGLLRGQIASDHFPVGFIPGRGARR"
     gene            1729. .2301
                     /gene="cdtC"
     CDS             1729. .2301
                     /gene="cdtC"
                     /experiment="experimental evidence, no additional details
                     recorded"
                     /codon_start=1
                     /transl_table=11
                     /product="CdtC"
                     /protein_id="AAD10623.1"
                     /db_xref="GI:416216"
                     /translation="MKTVIVFFVLLLTGCASEPANQRNLLTQFVGNNAPVDPEPSPVL
                     VNIRNVLTGGIIRNPVGSDFNVNNWVISEVKTNDLDLISAPGGHVQIKNPDGNECFAI
                     LNGQLAVAKQCSESDRNALFTFITSDTGAVQIKSIGSGQCLGNGESITDFRLKKCVDD
                     LGRPFDTVPPGLLWMLNPPLSPAIMSPLTS"
ORIGIN

  Query Match             99.8%;  Score 467;  DB 15;  Length 2305;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-291;
  Matches  467;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 GAAAATAAATGGAACACACATGTCCGACAACTTGTTACGGGAAGTGGTGCTGTTGATATC 61
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1019 GAAAATAAATGGAACACACATGTCCGACAACTTGTTACGGGAAGTGGTGCTGTTGATATC 1078

Qy         62 CTGATGGTTCAGGAGGCAGGGGCAGTACCAGCTTCTGCAACGTTGACTGAGCGAGAATTT 121
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1079 CTGATGGTTCAGGAGGCAGGGGCAGTACCAGCTTCTGCAACGTTGACTGAGCGAGAATTT 1138

Qy        122 AGCACTCCTGGTATTCCGATGAATGAGTATATCTGGAATACCGGAACCAATAGTCGCCCA 181
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1139 AGCACTCCTGGTATTCCGATGAATGAGTATATCTGGAATACCGGAACCAATAGTCGCCCA 1198

Qy        182 CAGGAGTTGTTTATATATTTCTCACGTGTTGATGCATTCGCTAACAGAGTAAATCTTGCG 241
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1199 CAGGAGTTGTTTATATATTTCTCACGTGTTGATGCATTCGCTAACAGAGTAAATCTTGCG 1258

Qy        242 ATTGTTTCAAACAGAAGAGCTGATGAGGTGATTGTATTACCTCCTCCAACTGTTGTATCA 301
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1259 ATTGTTTCAAACAGAAGAGCTGATGAGGTGATTGTATTACCTCCTCCAACTGTTGTATCA 1318

Qy        302 CGACCGATCATCGGCATTAGAATTGGTAATGATGTTTTCTTCTCAACCCATGCATTGGCG 361
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1319 CGACCGATCATCGGCATTAGAATTGGTAATGATGTTTTCTTCTCAACCCATGCATTGGCG 1378

Qy        362 AATCGGGGCGTGGATTCAGGAGCAATTGTAAATAGTGTTTTTGAGTTCTTCAACAGACAA 421
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1379 AATCGGGGCGTGGATTCAGGAGCAATTGTAAATAGTGTTTTTGAGTTCTTCAACAGACAA 1438

Qy        422 ACGGATCCTATAAGACAGGCCGCTAACTGGATGATTGCAGGAGATTT 468
              |||||||||||||||||||||||||||||||||||||||||||||||
Db       1439 ACGGATCCTATAAGACAGGCCGCTAACTGGATGATTGCAGGAGATTT 1485


RESULT 3
AY351905
LOCUS       AY351905                462 bp    DNA     linear   BCT 07-NOV-2003
DEFINITION  Escherichia coli strain VTE1456 cytolethal distending toxin (cdtB)
            gene, partial cds.
ACCESSION   AY351905
VERSION     AY351905.1  GI:34101070
KEYWORDS    .
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1  (bases 1 to 462)
  AUTHORS   Pandey,M., Khan,A., Das,S.C., Sarkar,B., Kahali,S., Chakraborty,S.,
            Chattopadhyay,S., Yamasaki,S., Takeda,Y., Nair,G.B. and
            Ramamurthy,T.
  TITLE     Association of Cytolethal Distending Toxin Locus cdtB with
            Enteropathogenic Escherichia coli Isolated from Patients with Acute
            Diarrhea in Calcutta, India
  JOURNAL   J. Clin. Microbiol. 41 (11), 5277-5281 (2003)
   PUBMED   14605183
REFERENCE   2  (bases 1 to 462)
  AUTHORS   Pandey,M., Khan,A., Das,S.C., Sarkar,B., Kahali,S., Chakraborty,S.,
            Chattopadhyay,S., Yamasaki,S., Takeda,Y., Nair,G.B. and
            Ramamurthy,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUL-2003) Microbiology, National Institute of Cholera
            and Enteric Diseases, P-33, CIT Road, Scheme, XM, Beliaghata,
            Calcutta, West Bengal 700010, India
FEATURES             Location/Qualifiers
     source          1. .462
                     /organism="Escherichia coli"
                     /mol_type="genomic DNA"
                     /strain="VTE1456"
                     /db_xref="taxon:562"
     gene            <1. .>462
                     /gene="cdtB"
     CDS             <1. .>462
                     /gene="cdtB"
                     /codon_start=1
                     /transl_table=11
                     /product="cytolethal distending toxin"
                     /protein_id="AAQ57596.1"
                     /db_xref="GI:34101071"
                     /translation="ENKWNTHVRQLVTGSGAVDILMVQEAGAVPASATLTEREFSTPG
                     IPMNEYIWNTGTNSRPQELFIYFSRVDAFANRVNLAIVSNRRADEVIVLPPPTVVSRP
                     IIGIRIGNDVFFSTHALANRGVDSGAIVNSVFEFFNRQTDPIRQAANWMIAG"
ORIGIN

  Query Match             98.7%;  Score 462;  DB 15;  Length 462;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-288;
  Matches  462;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 GAAAATAAATGGAACACACATGTCCGACAACTTGTTACGGGAAGTGGTGCTGTTGATATC 61
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GAAAATAAATGGAACACACATGTCCGACAACTTGTTACGGGAAGTGGTGCTGTTGATATC 60

Qy         62 CTGATGGTTCAGGAGGCAGGGGCAGTACCAGCTTCTGCAACGTTGACTGAGCGAGAATTT 121
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGATGGTTCAGGAGGCAGGGGCAGTACCAGCTTCTGCAACGTTGACTGAGCGAGAATTT 120

Qy        122 AGCACTCCTGGTATTCCGATGAATGAGTATATCTGGAATACCGGAACCAATAGTCGCCCA 181
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGCACTCCTGGTATTCCGATGAATGAGTATATCTGGAATACCGGAACCAATAGTCGCCCA 180

Qy        182 CAGGAGTTGTTTATATATTTCTCACGTGTTGATGCATTCGCTAACAGAGTAAATCTTGCG 241
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGGAGTTGTTTATATATTTCTCACGTGTTGATGCATTCGCTAACAGAGTAAATCTTGCG 240

Qy        242 ATTGTTTCAAACAGAAGAGCTGATGAGGTGATTGTATTACCTCCTCCAACTGTTGTATCA 301
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATTGTTTCAAACAGAAGAGCTGATGAGGTGATTGTATTACCTCCTCCAACTGTTGTATCA 300

Qy        302 CGACCGATCATCGGCATTAGAATTGGTAATGATGTTTTCTTCTCAACCCATGCATTGGCG 361
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CGACCGATCATCGGCATTAGAATTGGTAATGATGTTTTCTTCTCAACCCATGCATTGGCG 360

Qy        362 AATCGGGGCGTGGATTCAGGAGCAATTGTAAATAGTGTTTTTGAGTTCTTCAACAGACAA 421
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AATCGGGGCGTGGATTCAGGAGCAATTGTAAATAGTGTTTTTGAGTTCTTCAACAGACAA 420

Qy        422 ACGGATCCTATAAGACAGGCCGCTAACTGGATGATTGCAGGA 463
              |||||||||||||||||||||||||||||||||||||||||||
Db        421 ACGGATCCTATAAGACAGGCCGCTAACTGGATGATTGCAGGA 462


RESULT 4
AY351906
LOCUS       AY351906                462 bp    DNA     linear   BCT 07-NOV-2003
```

```
DEFINITION  Escherichia coli strain GB1807 cytolethal distending toxin (cdtB)
            gene, partial cds.
ACCESSION   AY351906
VERSION     AY351906.1  GI:34101072
KEYWORDS    .
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1  (bases 1 to 462)
  AUTHORS   Pandey,M., Khan,A., Das,S.C., Sarkar,B., Kahali,S., Chakraborty,S.,
            Chattopadhyay,S., Yamasaki,S., Takeda,Y., Nair,G.B. and
            Ramamurthy,T.
  TITLE     Association of Cytolethal Distending Toxin Locus cdtB with
            Enteropathogenic Escherichia coli Isolated from Patients with Acute
            Diarrhea in Calcutta, India
  JOURNAL   J. Clin. Microbiol. 41 (11), 5277-5281 (2003)
   PUBMED   14605183
REFERENCE   2  (bases 1 to 462)
  AUTHORS   Pandey,M., Khan,A., Das,S.C., Sarkar,B., Kahali,S., Chakraborty,S.,
            Chattopadhyay,S., Yamasaki,S., Takeda,Y., Nair,G.B. and
            Ramamurthy,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-JUL-2003) Microbiology, National Institute of Cholera
            and Enteric Diseases, P-33, CIT Road, Scheme, XM, Beliaghata,
            Calcutta, West Bengal 700010, India
FEATURES             Location/Qualifiers
     source          1. .462
                     /organism="Escherichia coli"
                     /mol_type="genomic DNA"
                     /strain="GB1807"
                     /db_xref="taxon:562"
     gene            <1. .>462
                     /gene="cdtB"
     CDS             <1. .>462
                     /gene="cdtB"
                     /codon_start=1
                     /transl_table=11
                     /product="cytolethal distending toxin"
                     /protein_id="AAQ57597.1"
                     /db_xref="GI:34101073"
                     /translation="ENKWNTHVRQLVTGSGAVDILMVQEAGAVPASATLTEREFSTPG
                     IPMNEYIWNTGTNSRPQELFIYFSRVDAFANRVNLAIVSNRRADEVIVLPPPTVVSRP
                     IIGIRIGNDVFFSTHALANRGVDSGAIVNSVFEFFNRQTDPIRQAANWMIAG"
ORIGIN

  Query Match           98.7%;  Score 462;  DB 15;  Length 462;
  Best Local Similarity  100.0%;  Pred. No. 2.5e-288;
  Matches  462;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 GAAAATAAATGGAACACACATGTCCGACAACTTGTTACGGGAAGTGGTGCTGTTGATATC 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GAAAATAAATGGAACACACATGTCCGACAACTTGTTACGGGAAGTGGTGCTGTTGATATC 60

Qy         62 CTGATGGTTCAGGAGGCAGGGGCAGTACCAGCTTCTGCAACGTTGACTGAGCGAGAATTT 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGATGGTTCAGGAGGCAGGGGCAGTACCAGCTTCTGCAACGTTGACTGAGCGAGAATTT 120

Qy        122 AGCACTCCTGGTATTCCGATGAATGAGTATATCTGGAATACCGGAACCAATAGTCGCCCA 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGCACTCCTGGTATTCCGATGAATGAGTATATCTGGAATACCGGAACCAATAGTCGCCCA 180

Qy        182 CAGGAGTTGTTTATATATTTCTCACGTGTTGATGCATTCGCTAACAGAGTAAATCTTGCG 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGGAGTTGTTTATATATTTCTCACGTGTTGATGCATTCGCTAACAGAGTAAATCTTGCG 240

Qy        242 ATTGTTTCAAACAGAAGAGCTGATGAGGTGATTGTATTACCTCCTCCAACTGTTGTATCA 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATTGTTTCAAACAGAAGAGCTGATGAGGTGATTGTATTACCTCCTCCAACTGTTGTATCA 300

Qy        302 CGACCGATCATCGGCATTAGAATTGGTAATGATGTTTTCTTCTCAACCCATGCATTGGCG 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CGACCGATCATCGGCATTAGAATTGGTAATGATGTTTTCTTCTCAACCCATGCATTGGCG 360

Qy        362 AATCGGGGCGTGGATTCAGGAGCAATTGTAAATAGTGTTTTTGAGTTCTTCAACAGACAA 421
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AATCGGGGCGTGGATTCAGGAGCAATTGTAAATAGTGTTTTTGAGTTCTTCAACAGACAA 420

Qy        422 ACGGATCCTATAAGACAGGCCGCTAACTGGATGATTGCAGGA 463
              |||||||||||||||||||||||||||||||||||||||||||
Db        421 ACGGATCCTATAAGACAGGCCGCTAACTGGATGATTGCAGGA 462


RESULT 5
AY351907
LOCUS       AY351907                462 bp    DNA     linear   BCT 07-NOV-2003
DEFINITION  Escherichia coli strain GB1371 cytolethal distending toxin (cdtB)
            gene, partial cds.
ACCESSION   AY351907
VERSION     AY351907.1  GI:34101074
KEYWORDS    .
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1  (bases 1 to 462)
  AUTHORS   Pandey,M., Khan,A., Das,S.C., Sarkar,B., Kahali,S., Chakraborty,S.,
            Chattopadhyay,S., Yamasaki,S., Takeda,Y., Nair,G.B. and
            Ramamurthy,T.
  TITLE     Association of Cytolethal Distending Toxin Locus cdtB with
            Enteropathogenic Escherichia coli Isolated from Patients with Acute
            Diarrhea in Calcutta, India
  JOURNAL   J. Clin. Microbiol. 41 (11), 5277-5281 (2003)
   PUBMED   14605183
REFERENCE   2  (bases 1 to 462)
  AUTHORS   Pandey,M., Khan,A., Das,S.C., Sarkar,B., Kahali,S., Chakraborty,S.,
            Chattopadhyay,S., Yamasaki,S., Takeda,Y., Nair,G.B. and
            Ramamurthy,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JUL-2003) Microbiology, National Institute of Cholera
            and Enteric Diseases, P-33, CIT Road, Scheme, XM, Beliaghata,
            Calcutta, West Bengal 700010, India
FEATURES             Location/Qualifiers
     source          1. .462
                     /organism="Escherichia coli"
                     /mol_type="genomic DNA"
                     /strain="GB1371"
                     /db_xref="taxon:562"
     gene            <1. .>462
                     /gene="cdtB"
     CDS             <1. .>462
                     /gene="cdtB"
                     /codon_start=1
                     /transl_table=11
                     /product="cytolethal distending toxin"
                     /protein_id="AAQ57598.1"
                     /db_xref="GI:34101075"
                     /translation="ENKWNTHVRQLVTGSGAVDILMVQEAGAVPASATLTEREFSTPG
                     IPMNEYIWNTGTNSRPQELFIYFSRVDAFANRVNLAIVSNRRADEVIVLPPPTVVSRP
                     IIGIRIGNDVFFSTHALANRGVDSGAIVNSVFEFFNRQTDPIRQAANWMIAG"
ORIGIN

  Query Match           98.7%;  Score 462;  DB 15;  Length 462;
  Best Local Similarity  100.0%;  Pred. No. 2.5e-288;
  Matches  462;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 GAAAATAAATGGAACACACATGTCCGACAACTTGTTACGGGAAGTGGTGCTGTTGATATC 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GAAAATAAATGGAACACACATGTCCGACAACTTGTTACGGGAAGTGGTGCTGTTGATATC 60

Qy         62 CTGATGGTTCAGGAGGCAGGGGCAGTACCAGCTTCTGCAACGTTGACTGAGCGAGAATTT 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGATGGTTCAGGAGGCAGGGGCAGTACCAGCTTCTGCAACGTTGACTGAGCGAGAATTT 120

Qy        122 AGCACTCCTGGTATTCCGATGAATGAGTATATCTGGAATACCGGAACCAATAGTCGCCCA 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        121 AGCACTCCTGGTATTCCGATGAATGAGTATATCTGGAATACCGGAACCAATAGTCGCCCA 180

Qy        182 CAGGAGTTGTTTATATATTTCTCACGTGTTGATGCATTCGCTAACAGAGTAAATCTTGCG 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGGAGTTGTTTATATATTTCTCACGTGTTGATGCATTCGCTAACAGAGTAAATCTTGCG 240

Qy        242 ATTGTTTCAAACAGAAGAGCTGATGAGGTGATTGTATTACCTCCTCCAACTGTTGTATCA 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATTGTTTCAAACAGAAGAGCTGATGAGGTGATTGTATTACCTCCTCCAACTGTTGTATCA 300

Qy        302 CGACCGATCATCGGCATTAGAATTGGTAATGATGTTTTCTTCTCAACCCATGCATTGGCG 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CGACCGATCATCGGCATTAGAATTGGTAATGATGTTTTCTTCTCAACCCATGCATTGGCG 360

Qy        362 AATCGGGGCGTGGATTCAGGAGCAATTGTAAATAGTGTTTTTGAGTTCTTCAACAGACAA 421
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AATCGGGGCGTGGATTCAGGAGCAATTGTAAATAGTGTTTTTGAGTTCTTCAACAGACAA 420

Qy        422 ACGGATCCTATAAGACAGGCCGCTAACTGGATGATTGCAGGA 463
              |||||||||||||||||||||||||||||||||||||||||||
Db        421 ACGGATCCTATAAGACAGGCCGCTAACTGGATGATTGCAGGA 462
```

RESULT 6
AY426341
LOCUS AY426341 494 bp DNA linear BCT 27-JAN-2004
DEFINITION Escherichia coli 3015-99 cytolethal distending toxin B subunit (cdtB) gene, partial cds.
ACCESSION AY426341
VERSION AY426341.1 GI:37575478
KEYWORDS .
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.
REFERENCE 1 (bases 1 to 494)
AUTHORS Pickett,C.L., Lee,R.B., Eyigor,A., Elitzur,B., Fox,E.M. and Strockbine,N.A.
TITLE Patterns of Variations in Escherichia coli Strains That Produce Cytolethal Distending Toxin
JOURNAL Infect. Immun. 72 (2), 684-690 (2004)
PUBMED 14742509
REFERENCE 2 (bases 1 to 494)
AUTHORS Pickett,C.L., Lee,R.B., Eyigor,A., Elitzur,B., Fox,E.M. and Strockbine,N.A.
TITLE Direct Submission
JOURNAL Submitted (01-OCT-2003) Microbiology, Immunology, and Molecular Genetics, University of Kentucky, 800 Rose Street, Lexington, KY 40536, USA
FEATURES Location/Qualifiers
source 1. .494
/organism="Escherichia coli"
/mol_type="genomic DNA"
/strain="3015-99"
/db_xref="taxon:562"
gene <1. .>494
/gene="cdtB"
CDS <1. .>494
/gene="cdtB"
/note="nuclease"
/codon_start=1
/transl_table=11
/product="cytolethal distending toxin B subunit"
/protein_id="AAQ93803.1"
/db_xref="GI:37575479"
/translation="VATWNLQGSNAPTENKWNTHVRQLVTGSGAVDILMVQEAGAVPA
SATLTEREFSTPGIPMNEYIWNTGTNSRPQELFIYFSRVDAFANRVNLAIVSNRRADE
VIVLPPPTVVSRPIIGIRIGNDVFFSTHALANRGVDSGAIVNSVFEFFNRQTDPIRQA
ANWM"
ORIGIN

Query Match 64.5%; Score 302; DB 15; Length 494;
Best Local Similarity 99.3%; Pred. No. 4.1e-184;

Matches 452; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          2 GAAAATAAATGGAACACACATGTCCGACAACTTGTTACGGGAAGTGGTGCTGTTGATATC 61
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         40 GAAAATAAATGGAACACACATGTCCGACAACTTGTTACGGGAAGTGGTGCTGTTGATATC 99

Qy         62 CTGATGGTTCAGGAGGCAGGGGCAGTACCAGCTTCTGCAACGTTGACTGAGCGAGAATTT 121
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        100 CTGATGGTTCAGGAGGCAGGGGCAGTACCAGCTTCTGCAACGTTGACTGAGCGAGAATTT 159

Qy        122 AGCACTCCTGGTATTCCGATGAATGAGTATATCTGGAATACCGGAACCAATAGTCGCCCA 181
              |||||||| |||||||||||||||||||||||||||||||||||||||||||| ||| |||
Db        160 AGCACTCCCGGTATTCCGATGAATGAGTATATCTGGAATACCGGAACCAATAGTCGTCCA 219

Qy        182 CAGGAGTTGTTTATATATTTCTCACGTGTTGATGCATTCGCTAACAGAGTAAATCTTGCG 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 CAGGAGTTGTTTATATATTTCTCACGTGTTGATGCATTCGCTAACAGAGTAAATCTTGCG 279

Qy        242 ATTGTTTCAAACAGAAGAGCTGATGAGGTGATTGTATTACCTCCTCCAACTGTTGTATCA 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        280 ATTGTTTCAAACAGAAGAGCTGATGAGGTGATTGTATTACCTCCTCCAACTGTTGTATCA 339

Qy        302 CGACCGATCATCGGCATTAGAATTGGTAATGATGTTTTCTTCTCAACCCATGCATTGGCG 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        340 CGACCGATCATCGGCATTAGAATTGGTAATGATGTTTTCTTCTCAACCCATGCATTGGCG 399

Qy        362 AATCGGGGCGTGGATTCAGGAGCAATTGTAAATAGTGTTTTTGAGTTCTTCAACAGACAA 421
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        400 AATCGAGGCGTGGATTCAGGAGCAATTGTAAATAGTGTTTTTGAGTTCTTCAACAGACAA 459

Qy        422 ACGGATCCTATAAGACAGGCCGCTAACTGGATGAT 456
              |||||||||||||||||||||||||||||||||||
Db        460 ACGGATCCTATAAGACAGGCCGCTAACTGGATGAT 494
```

RESULT 7
AF373205
LOCUS AF373205 463 bp DNA linear BCT 22-JUL-2005
DEFINITION Escherichia coli strain MBU. E 322 cytolethal distending toxin B (cdtB) gene, partial cds.
ACCESSION AF373205
VERSION AF373205.1 GI:20385545
KEYWORDS .
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.
REFERENCE 1 (bases 1 to 463)
AUTHORS Bouzari,S., Oloomi,M. and Oswald,E.
TITLE Detection of the cytolethal distending toxin locus cdtB among diarrheagenic Escherichia coli isolates from humans in Iran
JOURNAL Res. Microbiol. 156 (2), 137-144 (2005)
PUBMED 15748977
REFERENCE 2 (bases 1 to 463)
AUTHORS Bouzari,S., Oloomi,M. and Zarepoor,M.
TITLE Direct Submission
JOURNAL Submitted (22-APR-2001) Molecular Biology, Institute Pasteur of Iran, Pasteur Ave., Tehran 13164, Iran
FEATURES Location/Qualifiers
source 1. .463
/organism="Escherichia coli"
/mol_type="genomic DNA"
/strain="MBU. E 322"
/db_xref="taxon:562"
gene <1. .>463
/gene="cdtB"
CDS <1. .>463
/gene="cdtB"
/note="CdtB"
/codon_start=2
/transl_table=11
/product="cytolethal distending toxin B"
/protein_id="AAM21323.1"

```
                     /db_xref="GI:20385546"
                     /translation="XNKWNTHVRLLVTGSGAVDVXMVQEAGAVPASATLTEREFSTPG
                     IPMNEYIWNTGTNSRPQELFIYFSRVDAFANRVNLAIVPNRRADEVIVLPPPTVVSRP
                     IIGIRIGNDVFFSTHALANRGVDSGAIVNSVFEFFNRQTDPIRQAANWMIAG"
ORIGIN

  Query Match           63.7%;  Score 298;  DB 15;  Length 463;
  Best Local Similarity  99.5%;  Pred. No. 1.7e-181;
  Matches  398;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         63 TGATGGTTCAGGAGGCAGGGGCAGTACCAGCTTCTGCAACGTTGACTGAGCGAGAATTTA 122
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 TGATGGTTCAGGAGGCAGGGGCAGTACCAGCTTCTGCAACGTTGACTGAGCGAGAATTTA 122

Qy        123 GCACTCCTGGTATTCCGATGAATGAGTATATCTGGAATACCGGAACCAATAGTCGCCCAC 182
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GCACTCCTGGTATTCCGATGAATGAGTATATCTGGAATACCGGAACCAATAGTCGCCCAC 182

Qy        183 AGGAGTTGTTTATATATTTCTCACGTGTTGATGCATTCGCTAACAGAGTAAATCTTGCGA 242
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 AGGAGTTGTTTATATATTTCTCACGTGTTGATGCATTCGCTAACAGAGTAAATCTTGCGA 242

Qy        243 TTGTTTCAAACAGAAGAGCTGATGAGGTGATTGTATTACCTCCTCCAACTGTTGTATCAC 302
               ||||| ||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        243 TTGTTCCAAACAGAAGAGCTGATGAGGTGATTGTATTACCTCCCCCAACTGTTGTATCAC 302

Qy        303 GACCGATCATCGGCATTAGAATTGGTAATGATGTTTTCTTCTCAACCCATGCATTGGCGA 362
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 GACCGATCATCGGCATTAGAATTGGTAATGATGTTTTCTTCTCAACCCATGCATTGGCGA 362

Qy        363 ATCGGGGCGTGGATTCAGGAGCAATTGTAAATAGTGTTTTTGAGTTCTTCAACAGACAAA 422
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 ATCGGGGCGTGGATTCAGGAGCAATTGTAAATAGTGTTTTTGAGTTCTTCAACAGACAAA 422

Qy        423 CGGATCCTATAAGACAGGCCGCTAACTGGATGATTGCAGG 462
               |||||||||||||||||||||||||||||||||||||||
Db        423 CGGATCCTATAAGACAGGCCGCTAACTGGATGATTGCAGG 462


RESULT 8
AY426340
LOCUS       AY426340                 495 bp    DNA     linear   BCT 27-JAN-2004
DEFINITION  Escherichia coli 3180-98 cytolethal distending toxin B subunit
            (cdtB) gene, partial cds.
ACCESSION   AY426340
VERSION     AY426340.1  GI:37575476
KEYWORDS    .
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1  (bases 1 to 495)
  AUTHORS   Pickett,C.L., Lee,R.B., Eyigor,A., Elitzur,B., Fox,E.M. and
            Strockbine,N.A.
  TITLE     Patterns of Variations in Escherichia coli Strains That Produce
            Cytolethal Distending Toxin
  JOURNAL   Infect. Immun. 72 (2), 684-690 (2004)
   PUBMED   14742509
REFERENCE   2  (bases 1 to 495)
  AUTHORS   Pickett,C.L., Lee,R.B., Eyigor,A., Elitzur,B., Fox,E.M. and
            Strockbine,N.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-OCT-2003) Microbiology, Immunology, and Molecular
            Genetics, University of Kentucky, 800 Rose Street, Lexington, KY
            40536, USA
FEATURES             Location/Qualifiers
     source          1. .495
                     /organism="Escherichia coli"
                     /mol_type="genomic DNA"
                     /strain="3180-98"
                     /db_xref="taxon:562"
     gene            <1. .>495
                     /gene="cdtB"
     CDS             <1. .>495
                     /gene="cdtB"
                     /note="nuclease"
                     /codon_start=1
                     /transl_table=11
                     /product="cytolethal distending toxin B subunit"
                     /protein_id="AAQ93802.1"
                     /db_xref="GI:37575477"
                     /translation="VATWNLQGSNAPTENKWNTHVRQLVTGSGAVDILMVQEAGAVPA
                     SATLTEREFSTPGIPMNEYIWNTGTNSRPQELFIYFSRVDAFANGVNLAIVSNRRADE
                     VIVLPPPTVVSRPIIGIRIGNDVFFSTHALANRGVDSGAIVNSVFEFFNRQTDPIRQA
                     ANWMI"
ORIGIN

  Query Match           39.5%;  Score 185;  DB 15;  Length 495;
  Best Local Similarity  99.0%;  Pred. No. 6.8e-108;
  Matches  385;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GAAAATAAATGGAACACACATGTCCGACAACTTGTTACGGGAAGTGGTGCTGTTGATATC 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         40 GAAAATAAATGGAACACACATGTCCGACAACTTGTTACGGGAAGTGGTGCTGTTGATATC 99

Qy         62 CTGATGGTTCAGGAGGCAGGGGCAGTACCAGCTTCTGCAACGTTGACTGAGCGAGAATTT 121
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        100 CTGATGGTTCAGGAGGCAGGGGCAGTACCAGCTTCTGCCACGTTGACTGAGCGAGAATTT 159

Qy        122 AGCACTCCTGGTATTCCGATGAATGAGTATATCTGGAATACCGGAACCAATAGTCGCCCA 181
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        160 AGCACTCCCGGTATTCCGATGAATGAGTATATCTGGAATACCGGAACCAATAGTCGTCCA 219

Qy        182 CAGGAGTTGTTTATATATTTCTCACGTGTTGATGCATTCGCTAACAGAGTAAATCTTGCG 241
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        220 CAGGAGTTGTTTATATATTTCTCACGTGTTGATGCATTCGCTAACGGAGTAAATCTTGCG 279

Qy        242 ATTGTTTCAAACAGAAGAGCTGATGAGGTGATTGTATTACCTCCTCCAACTGTTGTATCA 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        280 ATTGTTTCAAACAGAAGAGCTGATGAGGTGATTGTATTACCTCCTCCAACTGTTGTATCA 339

Qy        302 CGACCGATCATCGGCATTAGAATTGGTAATGATGTTTTCTTCTCAACCCATGCATTGGCG 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        340 CGACCGATCATCGGCATTAGAATTGGTAATGATGTTTTCTTCTCAACCCATGCATTGGCG 399

Qy        362 AATCGGGGCGTGGATTCAGGAGCAATTGT 390
              |||||||||||||||||||||||||||||
Db        400 AATCGGGGCGTGGATTCAGGAGCAATTGT 428


RESULT 9
AX140193
LOCUS       AX140193                 822 bp    DNA     linear   PAT 31-MAY-2001
DEFINITION  Sequence 3 from Patent WO0134205.
ACCESSION   AX140193
VERSION     AX140193.1  GI:14280533
KEYWORDS    .
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1
  AUTHORS   Purdy,D.M., Charlton,S.M. and Henderson,I.D.
  TITLE     Use of lytic toxins and toxin conjugates
  JOURNAL   Patent: WO 0134205-A 3 17-MAY-2001;
            Microbiological Research Authority (GB)
FEATURES             Location/Qualifiers
     source          1. .822
                     /organism="Escherichia coli"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:562"
ORIGIN

  Query Match           9.4%;  Score 44;  DB 2;  Length 822;
  Best Local Similarity  100.0%;  Pred. No. 4.9e-16;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy        425 GATCCTATAAGACAGGCCGCTAACTGGATGATTGCAGGAGATTT 468
              ||||||||||||||||||||||||||||||||||||||||||||
Db        532 GATCCTATAAGACAGGCCGCTAACTGGATGATTGCAGGAGATTT 575


RESULT 10
AY578329
LOCUS       AY578329               10241 bp    DNA     linear   BCT 01-APR-2005
DEFINITION  Escherichia coli strain 28C lambdoid prophage host specificity
            protein gene, partial cds; putative membrane protein precursor,
            putative tail fiber protein, hypothetical proteins, cytolethal
            distending toxin type IV subunit A (cdtA), cytolethal distending
            toxin type IV subunit B (cdtB), and cytolethal distending toxin
            type IV subunit C (cdtC) genes, complete cds; and hypothetical
            protein gene, partial cds.
ACCESSION   AY578329
VERSION     AY578329.1  GI:50983070
KEYWORDS    .
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1  (bases 1 to 10241)
  AUTHORS   Ledger,N., Pujiwara,T., Boury,M., Sugai,M. and Oswald,E.
  TITLE     Escherichia coli 28C (O75) cytolethal distending toxin type IV
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 10241)
  AUTHORS   Oswald,E.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-MAR-2004) UMR1225, INRA-ENVT, 23 Chemin des Capelles,
            Toulouse 31000, France
FEATURES             Location/Qualifiers
     source          1. .10241
                     /organism="Escherichia coli"
                     /mol_type="genomic DNA"
                     /strain="28C"
                     /db_xref="taxon:562"
                     /note="serogroup: O75; pathovar: ExPEC; endogenous_virus:
                     lambdoid prophage"
     CDS             <1. .633
                     /note="Orf1"
                     /codon_start=1
                     /transl_table=11
                     /product="host specificity protein"
                     /protein_id="AAT92041.1"
                     /db_xref="GI:50983071"
                     /translation="IFMNEVFLKYLTAPTITSGGNPPAFSLTPDGRLTAKNADISGNV
                     NANSGTLNNVTINQNCRILGKLSANQIEGDIVKTVGKAFPRNGSYASGTITVTVYDDQ
                     AFDRQIVIPPVLFRGGKHENFNSNNQQSYWYSTCKLQVLKNGQEIFQQPATDVSRVFS
                     SVIDMPAGHGHVTLTFNVSSYGANNWTPTTSISDLLVVVMKKSTAGITIS"
     CDS             702. .1202
                     /note="Orf2; similar to membrane protein Lom"
                     /codon_start=1
                     /transl_table=11
                     /product="putative membrane protein precursor"
                     /protein_id="AAT92042.1"
                     /db_xref="GI:50983072"
                     /translation="MRKLCAVILSAVVWLVAAGTPASAAEHQSTLSAGYLQTHTDMPG
                     SDDLKGINVKYRYEFTDTLGLVTSFSYTNAKDEQKTHYSDTRWHEDSVRNRWFSMMAG
                     PSVRVNEWFSAYAMAGVAYSRVSTFSGDYLRVTDNKGKTHDVLTGSDGGRHSNTVSGV
                     GGWRAV"
     CDS             1367. .3535
                     /note="Orf3; similar to ECs1650 from Escherichia coli
                     O157:H7 Sakai"
                     /codon_start=1
                     /transl_table=11
                     /product="putative tail fiber protein"
                     /protein_id="AAT92043.1"
                     /db_xref="GI:50983073"
                     /translation="MAAVKISGVLKDGAGKPIQNCTIQLKAKRNSTTVLVNTVASENP
                     DEAGRYSMDVEYGQYSVILLVEGFPPSHAGTITVYEGSRPGTLNDFLGAMTEEDVMPE
                     ALRRFEAMVEEVARNAEAASQSAAAAKKSETAAASSKNAAKTSETNAANSAQAAATSK
                     TASANSATAAKKSETNAKNSETAAKTSETNAKSSQTAAKTSETNAKASETAAKNSQAA
                     AAESESAAAGSATSAAGSATAAANSQKAAKTSETNAKSSQTAAKTSETNAKASETAAK
                     NSQDAAAQSESAAAGSATAAANSQKAAKTSETNAKTSETKAAASETEATSAATRAENA
                     ARIAEDAADPASVPPLPDIWLPLNDSLEAITGYAPGYKTITIGSDEITMPVNGICQFS
                     WASSATYIDKSGHITVAGNNVPRFEKYGLLIENQRTNMFVNSFNPDAWNKSGGISVTS
                     STDEFEFKYGRFTVGSDIAGTTTGRNICTVAGNKGIDVTGDDQYSKGPYVTASFRVRS
                     DLNVRARIRFERYNSEGYTFLCDAYLSLQTHELQITGGNAQLLTANFEIDPGSGWIYF
                     QATLKCLPEWGMVGTQLQIAADRAVGSFATGDWIEVTTPQFEYGACATSFIITTTEPA
                     TRASDLCKFPPLMKNMYTMPFTFMVEVHKNWFISHNAAPRVIDSENHQSGGPFIMGFGS
                     SGTISQDGYSYCDIGGANRRVYESCGVRDLVMDSGLRLTA"
     CDS             3723. .3992
                     /inference="non-experimental evidence, no additional
                     details recorded"
                     /note="Orf4; similar to c3149 from Escherichia coli CFT073
                     (serotype O6:H1:K?)"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAT92044.1"
                     /db_xref="GI:50983074"
                     /translation="MRDITLRFDNREQFNAIVYDSGLFSLEEENGVLVDVIGCIIDYE
                     EPENERCTGIDRGGFFVNMRIVDGSKNISSLMPFITTDQHVRTWA"
     CDS             4005. .4688
                     /inference="non-experimental evidence, no additional
                     details recorded"
                     /note="Orf5; similar to hypothetical protein in EAEC 042"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAT92045.1"
                     /db_xref="GI:50983075"
                     /translation="MVTKTVIPGDITTLKGDISSINGKVSTLQTEMTSAKQDISSRYT
                     KTEVDNKLKNKVEVNDLESGRYGGDFYPLTGREAFYLWNLATTTAAANLYLNPDPAIS
                     SVLRSTSSIRYKHSVETIDSEHAALIFRMRSVWYRSQCENDRRDWGFYGLIAEEVGEI
                     APQFVHWRPANEDDAPETISSNGLVAEGVMYERLVVPLIHHIQKLTERVDELESELKL
                     LLTSRSDIG"
     CDS             4699. .4989
                     /inference="non-experimental evidence, no additional
                     details recorded"
                     /note="Orf6; similar to c3147 from Escherichia coli CFT073
                     (serotype O6:H1:K?)"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAT92046.1"
                     /db_xref="GI:50983076"
                     /translation="MDITPFLHALCAVAAQVLVGLFTGNWAYGAIAGCTFFIAREHTQ
                     AEYRWIEMFGHGKRINMPWWGGFDLRAWDMASMMDFAVPVVACLLVWLLVNR"
     gene            5502. .6215
                     /gene="cdtA"
                     /note="cdtA-IV"
     CDS             5502. .6215
                     /gene="cdtA"
                     /note="CdtA-IV"
                     /codon_start=1
                     /transl_table=11
                     /product="cytolethal distending toxin type IV subunit A"
                     /protein_id="AAT92047.1"
                     /db_xref="GI:50983077"
                     /translation="MDKKLIAFLCTLIITGCSDGIGDSPSPPGKNVELVGIPGQGVAV
                     ASNGTSPTFGSNSTGFPDVSIMSTGGAMLTVWARPVRNWLWGYTPFDSVSFGENRNWK
                     VVDGKDAGTVKFVNVAQGTCMEAFKNGVIHNTCDDNSLSQEFQLLPSTNGNVLIRSSA
                     LQTCIRADYLSRTILSPFAFTITLEKCPGAKEETQEMLWAISPPVRAAKPNLIKPELR
                     PFRPLPIPPHDKPDGMEGV"
     gene            6212. .7033
                     /gene="cdtB"
                     /note="cdtB-IV"
     CDS             6212. .7033
                     /gene="cdtB"
                     /note="CdtB-IV"
                     /codon_start=1
                     /transl_table=11
                     /product="cytolethal distending toxin type IV subunit B"
```

```
                     /protein_id="AAT92048.1"
                     /db_xref="GI:50983078"
                     /translation="MKKLLFLLMILPGISFADLSDFKVATWNLQGSNAPTENKWNTHV
                     RQLVTGSGAVDILMVQEAGSIPSSATLTEREFRTPGIPMNEYIWNTGTNSRPQQLFIY
                     FSRTDALSNRVNLAIVSNRRADEVIVLSPPTVASRPIIGIRIGNDVFFSTHALANRGI
                     DSGAIVNSVFEFFNRQTDPIRQAANWMIAGDFNRSPAMLFSTLEPGIRNHVNIIAPPD
                     PTQASGGVLDYAVVGNSVSFVLPLLRASLLFGLLRGQIASDHFPVGFIPGRGARR"
     gene            7030. .7602
                     /gene="cdtC"
                     /note="cdtC-IV"
     CDS             7030. .7602
                     /gene="cdtC"
                     /note="CdtC-IV"
                     /codon_start=1
                     /transl_table=11
                     /product="cytolethal distending toxin type IV subunit C"
                     /protein_id="AAT92049.1"
                     /db_xref="GI:50983079"
                     /translation="MRIIITILVLLLTGCSSDVVNQRGLLTQFAGNNAPGDPEPRPVL
                     VSIRNVLTGGILRNPVGRDFNVNNWVISEVKTNDLDLISAPGGHVQIKNPDGNECLAI
                     LNGQLAVAKQCTESNRNALFTFITSETGAVQIKSIGNGQCLGNGESVTDFKLTKCVND
                     LSRPFDTVSPGLLWMLNPPLSPAIMSPLTI"
     CDS             complement(7672. .9156)
                     /inference="non-experimental evidence, no additional
                     details recorded"
                     /note="Rorf1; similar to ECs1662 from Escherichia coli
                     O157:H7 Sakai; contains a putative conserved domain
                     (alpha/beta hydrolase fold) found in a wide range of
                     enzymes"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAT92050.1"
                     /db_xref="GI:50983080"
                     /translation="MRKIIAHFKIVLTLLLPVTVSAQQIEWQSCMTSQFNHWFGEEKP
                     SPDLLCGYLSVPLKYTDTGGDASYEKKSQVKLALTKLPAKSKHKGSILIISGGPGLPG
                     INPYINFDWPVTNLRESWDIIGFDPRGVGQSTPTINCQQSDTEIQENITEKQRVLNKI
                     NACIHNTGAEVIRHIGSNEAVYDIDRIRQALGDKQLTAVAYSYGTQIAALYAERFPYN
                     VRSIVLDGVVDIDDLEDNFTWQLKQAQSYQETFDRFASWCARTKSCPLSSDRDKAITQ
                     FHELLLKLHHRPLIDSKGENISSDELISLTTDLLLWRSSWPTLATAIRQFSQGIVSNE
                     IETALSAPIASEEPSDALGVILCVDQGDEKLTPEERKSRKDALANAFPAINFDNGRSD
                     SPDFCELWPIHSDLNKTRLKNTVLPSGLLFVAHKYDPTTPWINARKMAEKFSSPLLTI
                     NGDGHTLALTGVNLCVDKAVVHHLLTPQKTEDISCPGNAEVDIQ"
     CDS             complement(9341. .>10241)
                     /inference="non-experimental evidence, no additional
                     details recorded"
                     /note="Rorf2; similar to outer membrane protease OMPT
                     (protease VII precursor)"
                     /codon_start=2
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAT92051.1"
                     /db_xref="GI:50983081"
                     /translation="SFASTETLSFTPDNINADISLGTLSGKTKERVYLAEEGGRKVSQ
                     LDWKFNNAAIIKGAINWDLMPQVSVGAAGWTTLGQKGGNMIDRDWQDPDKPGIWTDES
                     RHPDTRLNFANEFDLNIKGWLLNESNYRLGLMAGYQESRYSFTARGGSYIYSDEGFRD
                     DIGSIPNGERAIGYKQRFKMPYIGLTGSYRYEDFEFGGTFKYSGWVEASDNDEHYARV
                     KRITYRSNVKDQNYYSIAVNAGYYVTPNAKVYVEGAWNRVTNKKGDTSLYDHNDNTSE
                     YSKNGAGIENYNFITTAGLKYTF"
ORIGIN

  Query Match              9.4%;  Score 44;  DB 15;  Length 10241;
  Best Local Similarity   100.0%;  Pred. No. 5.9e-16;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        425 GATCCTATAAGACAGGCCGCTAACTGGATGATTGCAGGAGATTT 468
              |||||||||||||||||||||||||||||||||||||||||||
Db       6743 GATCCTATAAGACAGGCCGCTAACTGGATGATTGCAGGAGATTT 6786


RESULT 11
AY162217
LOCUS       AY162217                 417 bp    DNA     linear   BCT 05-SEP-2003
DEFINITION  Escherichia coli cytolethal distending toxin B type IV (cdtB-IV)
            gene, partial cds.
ACCESSION   AY162217
VERSION     AY162217.1  GI:34482170
KEYWORDS    .
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1  (bases 1 to 417)
  AUTHORS   Toth,I., Herault,F., Beutin,L. and Oswald,E.
  TITLE     Production of Cytolethal Distending Toxins by Pathogenic
            Escherichia coli Strains Isolated from Human and Animal Sources:
            Establishment of the Existence of a New cdt Variant (Type IV)
  JOURNAL   J. Clin. Microbiol. 41 (9), 4285-4291 (2003)
   PUBMED   12958258
REFERENCE   2  (bases 1 to 417)
  AUTHORS   Oswald,E.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-OCT-2002) UMR960 de Microbiologie Moleculaire,
            INRA-ENVT, 23 chemin des Capelles, Toulouse 31000, France
FEATURES             Location/Qualifiers
     source          1. .417
                     /organism="Escherichia coli"
                     /mol_type="genomic DNA"
                     /strain="28C"
                     /db_xref="taxon:562"
                     /note="serogroup: O75; pathovar: ExPEC"
     gene            <1. .>417
                     /gene="cdtB-IV"
     CDS             <1. .>417
                     /gene="cdtB-IV"
                     /codon_start=2
                     /transl_table=11
                     /product="cytolethal distending toxin B type IV"
                     /protein_id="AAO25510.1"
                     /db_xref="GI:34482171"
                     /translation="QLVTGSGAVDILMVQEAGSIPSSATLTEREFRTPGIPMNEYIWN
                     TGTNSRPQQLFIYFSRTDALSNRVNLAIVSNRRADEVIVLSPPTVASRPIIGIRIGND
                     VFFSTHALANRGIDSGAIVNSVFEFFNRQTDPIRQAA"
ORIGIN

  Query Match              8.8%;  Score 41;  DB 15;  Length 417;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-14;
  Matches   41;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         38 ACGGGAAGTGGTGCTGTTGATATCCTGATGGTTCAGGAGGC 78
              |||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGGAAGTGGTGCTGTTGATATCCTGATGGTTCAGGAGGC 51


RESULT 12
BX465849/c
LOCUS       BX465849               136195 bp    DNA     linear   VRT 10-NOV-2005
DEFINITION  Zebrafish DNA sequence from clone DKEY-263L8 in linkage group 1,
            complete sequence.
ACCESSION   BX465849
VERSION     BX465849.10  GI:81685375
KEYWORDS    HTG.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 136195)
  AUTHORS   Barker,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-NOV-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests:
            http://www.sanger.ac.uk/Projects/D_rerio/faqs.shtml#dataeight
COMMENT     On Nov 10, 2005 this sequence version replaced gi:41392538.
            ---------- Genome Center
```

Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
Clone-derived Zebrafish pUC subclones occasionally display inconsistency over the length of mononucleotide A/T runs and conserved TA repeats. Where this is found the longest good quality representation will be submitted.
Any regions longer than 1kb tagged as misc-feature 'unsure' are part of a tandem repeat of more than 10kb in length where it has not been possible to anchor the base differences between repeat copies. The region has been built up based on the repeat element to match the total size of repeat indicated by restriction digest, but repeat copies may not be in the correct order and the usual finishing criteria may not apply.
DKEY-263L8 is from a Zebrafish BAC library
VECTOR: pIndigoBAC-5.

FEATURES Location/Qualifiers
source 1. .136195
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/chromosome="1"
/clone="DKEY-263L8"
/clone_lib="DanioKey"

ORIGIN

Query Match 4.5%; Score 21; DB 11; Length 136195;
Best Local Similarity 100.0%; Pred. No. 0.68;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        400 TTTTGAGTTCTTCAACAGACA 420
              ||||||||||||||||||||
Db      95145 TTTTGAGTTCTTCAACAGACA 95125
```

RESULT 13
AC128885
LOCUS AC128885 141587 bp DNA linear HTG 20-NOV-2002
DEFINITION Rattus norvegicus clone CH230-334P14, *** SEQUENCING IN PROGRESS ***, 4 unordered pieces.
ACCESSION AC128885
VERSION AC128885.3 GI:25139458
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 141587)
AUTHORS Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 141587)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (24-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3 (bases 1 to 141587)
AUTHORS Rat Genome Sequencing Consortium.
TITLE Direct Submission
JOURNAL Submitted (20-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT On Nov 20, 2002 this sequence version replaced gi:23195910.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.
---------- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
---------- Project Information
Center project name: KAXT

```
           Center clone name: CH230-334P14
          ------------ Summary Statistics
           Assembly program: Phrap; version 0.990329
           Consensus quality: 127909 bases at least Q40
           Consensus quality: 130224 bases at least Q30
           Consensus quality: 131369 bases at least Q20
           Estimated insert size: 129726; sum-of-contigs estimation
           Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
          ------------
          * NOTE: Estimated insert size may differ from sequence length
          *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
          * NOTE: This is a 'working draft' sequence. It currently
          * consists of 4 contigs. The true order of the pieces
          * is not known and their order in this sequence record is
          * arbitrary. Gaps between the contigs are represented as
          * runs of N, but the exact sizes of the gaps are unknown.
          * This record will be updated with the finished sequence
          * as soon as it is available and the accession number will
          * be preserved.
          *        1   105419: contig of 105419 bp in length
          *   105420   105519: gap of unknown length
          *   105520   138656: contig of 33137 bp in length
          *   138657   138756: gap of unknown length
          *   138757   140163: contig of 1407 bp in length
          *   140164   140263: gap of unknown length
          *   140264   141587: contig of 1324 bp in length.
FEATURES             Location/Qualifiers
     source          1. .141587
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-334P14"
     gap             105420. .105519
                     /estimated_length=unknown
     misc_feature    105520. .107531
                     /note="wgs_contig"
     gap             138657. .138756
                     /estimated_length=unknown
     gap             140164. .140263
                     /estimated_length=unknown
ORIGIN

  Query Match              4.5%;  Score 21;  DB 12;  Length 141587;
  Best Local Similarity   100.0%;  Pred. No. 0.68;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        388 TGTAAATAGTGTTTTTGAGTT 408
              |||||||||||||||||||||
Db      84315 TGTAAATAGTGTTTTTGAGTT 84335


RESULT 14
AC175073/c
LOCUS       AC175073               191669 bp    DNA     linear   HTG 31-JAN-2006
DEFINITION  Bos taurus clone CH240-245G12, *** SEQUENCING IN PROGRESS ***, 29
            unordered pieces.
ACCESSION   AC175073
VERSION     AC175073.2  GI:86200618
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Bos taurus (cattle)
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
            Pecora; Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 191669)
  AUTHORS   Muzny,D., Adams,C., Agbai II,O., Allen,C., Alsbrooks,S., Archer,P.,
            Arredondo,H., Bandaranaike,D., Bangura,L., Beltran,B., Beltran,R.,
            Beraducci,A., Biswalo,K., Blyth,P., Bonham,H., Buhay,C., Burch,P.,
            Cadoree,I., Canada,A., Cardenas,V., Carter,K., Cavazos,I.,
            Chacko,J., Chahrour,M., Chavez,D., Chen,A., Chen,G., Chen,R.,
            Cheng,M.-T., Chu,J., Clerc,K., Cockrell,R., Coyle,M., Cree,A.,
            Curry,S., Dai,W., Davila,M.L., Davis,C., Davy-Carroll,L., De
            Anda,C., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H.,
            Donlin,J., McCauley,S., Dugan-Rocha,S., Dunn,A., Durbin,K.,
            Dziuda,D., Egan,A., Escotto,M., Espinosa,V., Eugene,C., Fa,M.,
            Fernandez,S., Fernando,P., Flagg,N., Forbes,L., Foster,P.,
            Fowler,G., Fu,Q., Fuh,E., Garcia,A., Garcia,R., Garner,T.,
            Gaskin,C., Gench,S., Ghose,S., Gill,R., Gonzalez,D.,
            Gonzalez-Garay,M., Guevara,W., Holder,M., Haaland,W., Haeberlen,K.,
            Hall,B., Hamid,H., Hamilton,K., Harbes,B., Harris,R., Havlak,P.,
            Hawes,A., Hawkins,E., Hayes,S., Hemphill,L., Hernandez,J.,
            Hines,S., Hitchens,M., Hodgson,A., Hogues,M., Hollins,B.,
            Howell,L.T., Hulyk,S., Hume,J., Imo,K., Jackson,A., Jackson,L.,
            Jacob,L., Jiang,H., Johnson,B., Johnson,R., Kalafus,K., Kelly,S.,
            Keys,T., Khan,Z., King,L., Kovar,C., Kowis,A., Kowis,C., Lara,F.,
            Leal,S., Lee,K., Lee,S., LeGall,F.I., Lemon,S., Lewis,L., Li,B.,
            Li,Y., Li,Z., Linnell,M., Liu,W., Liu,Y.-S., Liu,Y., Liyanage,D.,
            London,P., Lopez,J., Lorensuhewa,L., Lozado,R., Luk,T., Madu,R.,
            Maheshwari,M., Mahoney,C., Malloy,K., Mansouri,D., Martinez,E.,
            McClelland,H., McPherson,J., Mercadao,C., Metzker,M.,
            Milosavljevic,A., Minja,E., Morgan,M., Morris,S., Munidasa,M.,
            Murray,D., Nazarith,L., Ngo,D., Nguyen,N., Norwig-Eastaugh,E.,
            Nott,A., Nwaokelemeh,O., Obregon,M., Ochi-Okorie,C., Odeh,E.,
            Okwuonu,G., Okwuonu,K., Parker,D., Pasternak,S., Patel,B.,
            Patel,V., Paul,H., Perez,A., Perez,L., Petrosino,J., Pham,T.,
            Primus,E., Pu,L.-L., Puazo,M., Qin,X., Quinn,A., Quiroz,J.,
            Rabata,D., Rachlin,E., Reigh,R., Ren,Y., Reuter,M., Richards,S.,
            Rives,C., Rodriguez,F., Rojas,A., Ruiz,S.J., Sana,M., Sanders,W.,
            Santibanez,J., Santos,R., Savery,G., Scherer,S., Shen,H., Shen,Y.,
            Sisson,I., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R.,
            Svatek,A., Taylor,E., Taylor,T., Thomas,N., Thorn,R., Thornton,R.,
            Trejos,Z., Usmani,K., Vargo,C., Verduzco,D., Villasana,D., Virk,D.,
            Volkov,A., Waldron,L., Walker,B., Wang,Q., Wang,S., Warren,J.,
            Wei,X., Wheeler,D., Williams,G., Williams,R., Worley,K., Wright,R.,
            Wu,J., Yakub,S., Yan,K., Yuan,Y., Yu,F., Zhang,J., Zhang,L.,
            Zhang,Z., Zhou,J., Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 191669)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-JAN-2006) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 191669)
  AUTHORS   .
  CONSRTM   Bovine Genome Sequencing Consortium
  TITLE     Direct Submission
  JOURNAL   Submitted (31-JAN-2006) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Jan 31, 2006 this sequence version replaced gi:84490450.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: FNEW
                Center clone name: CH240-245G12
            -------------- Summary Statistics
                Assembly program: Atlas 3.0;
                Consensus quality: 168512 bases at least Q40
                Consensus quality: 171717 bases at least Q30
```

```
    Consensus quality: 173988 bases at least Q20
    Estimated insert size: 175814; sum-of-contigs estimation
    Quality coverage: 5x in Q20 bases; sum-of-contigs estimation
--------------
* NOTE: Estimated insert size may differ from sequence length
*   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 29 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*        1     1361: contig of 1361 bp in length
*     1362     2075: gap of 714 bp
*     2076     6680: contig of 4605 bp in length
*     6681     6780: gap of unknown length
*     6781    11436: contig of 4656 bp in length
*    11437    12103: gap of 667 bp
*    12104    13463: contig of 1360 bp in length
*    13464    13670: gap of 207 bp
*    13671    17094: contig of 3424 bp in length
*    17095    17503: gap of 409 bp
*    17504    21392: contig of 3889 bp in length
*    21393    21492: gap of unknown length
*    21493    36367: contig of 14875 bp in length
*    36368    36417: gap of 50 bp
*    36418    42996: contig of 6579 bp in length
*    42997    43046: gap of 50 bp
*    43047    55330: contig of 12284 bp in length
*    55331    55380: gap of 50 bp
*    55381    57120: contig of 1740 bp in length
*    57121    65692: gap of 8572 bp
*    65693    68984: contig of 3292 bp in length
*    68985    69034: gap of 50 bp
*    69035    71002: contig of 1968 bp in length
*    71003    71052: gap of 50 bp
*    71053    77566: contig of 6514 bp in length
*    77567    78438: gap of 872 bp
*    78439    82513: contig of 4075 bp in length
*    82514    82563: gap of 50 bp
*    82564    92823: contig of 10260 bp in length
*    92824    92873: gap of 50 bp
*    92874    96190: contig of 3317 bp in length
*    96191    96240: gap of 50 bp
*    96241    99785: contig of 3545 bp in length
*    99786   100387: gap of 602 bp
*   100388   108057: contig of 7670 bp in length
*   108058   108107: gap of 50 bp
*   108108   120856: contig of 12749 bp in length
*   120857   120906: gap of 50 bp
*   120907   126815: contig of 5909 bp in length
*   126816   126865: gap of 50 bp
*   126866   133104: contig of 6239 bp in length
*   133105   133204: gap of unknown length
*   133205   161407: contig of 28203 bp in length
*   161408   161457: gap of 50 bp
*   161458   178544: contig of 17087 bp in length
*   178545   178644: gap of unknown length
*   178645   180645: contig of 2001 bp in length
*   180646   180745: gap of unknown length
*   180746   181866: contig of 1121 bp in length
*   181867   181966: gap of unknown length
*   181967   183136: contig of 1170 bp in length
*   183137   183236: gap of unknown length
*   183237   184890: contig of 1654 bp in length
*   184891   184990: gap of unknown length
*   184991   187827: contig of 2837 bp in length
*   187828   187927: gap of unknown length
*   187928   191669: contig of 3742 bp in length.
FEATURES             Location/Qualifiers
     source          1. .191669
                     /organism="Bos taurus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9913"
                     /clone="CH240-245G12"
          gap        1362. .2075
                     /estimated_length=714
          gap        6681. .6780
                     /estimated_length=unknown
          gap        11437. .12103
                     /estimated_length=667
          gap        13464. .13670
                     /estimated_length=207
          gap        17095. .17503
                     /estimated_length=409
          gap        21393. .21492
                     /estimated_length=unknown
          gap        36368. .36417
                     /estimated_length=50
          gap        42997. .43046
                     /estimated_length=50
          gap        55331. .55380
                     /estimated_length=50
          gap        57121. .65692
                     /estimated_length=8572
          gap        68985. .69034
                     /estimated_length=50
          gap        71003. .71052
                     /estimated_length=50
          gap        77567. .78438
                     /estimated_length=872
          gap        82514. .82563
                     /estimated_length=50
          gap        92824. .92873
                     /estimated_length=50

  Query Match             4.5%;  Score 21;  DB 12;  Length 191669;
  Best Local Similarity   100.0%;  Pred. No. 0.7;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        385 AATTGTAAATAGTGTTTTTGA 405
              |||||||||||||||||||||
Db      29658 AATTGTAAATAGTGTTTTTGA 29638


RESULT 15
AC107439/c
LOCUS       AC107439               238377 bp    DNA     linear   HTG 19-NOV-2002
DEFINITION  Rattus norvegicus clone CH230-229M1, *** SEQUENCING IN PROGRESS
            ***, 8 unordered pieces.
ACCESSION   AC107439
VERSION     AC107439.5  GI:25075945
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 238377)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
```

Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.

TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 238377)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (21-JAN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE 3 (bases 1 to 238377)
AUTHORS Rat Genome Sequencing Consortium.
TITLE Direct Submission
JOURNAL Submitted (19-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT On Nov 19, 2002 this sequence version replaced gi:23603151.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.
---------- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
---------- Project Information
Center project name: GNII
Center clone name: CH230-229M1
---------- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 176033 bases at least Q40
Consensus quality: 178158 bases at least Q30
Consensus quality: 179520 bases at least Q20
Estimated insert size: 178417; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 10410: contig of 10410 bp in length
* 10411 10510: gap of unknown length
* 10511 20957: contig of 10447 bp in length
* 20958 21057: gap of unknown length
* 21058 224179: contig of 203122 bp in length
* 224180 224279: gap of unknown length
* 224280 229532: contig of 5253 bp in length
* 229533 229632: gap of unknown length
* 229633 230665: contig of 1033 bp in length
* 230666 230765: gap of unknown length
* 230766 231997: contig of 1232 bp in length
* 231998 232097: gap of unknown length
* 232098 234781: contig of 2684 bp in length
* 234782 234881: gap of unknown length
* 234882 238377: contig of 3496 bp in length.

FEATURES Location/Qualifiers
source 1. .238377
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-229M1"
misc_feature 1. .1567
/note="wgs_contig"
gap 10411. .10510
/estimated_length=unknown
misc_feature 10511. .11540
/note="wgs_contig"
gap 20958. .21057
/estimated_length=unknown
misc_feature complement(200261. .201154)
/note="clone_boundary
clone_end:Sp6
site:
end_sequence:BZ108180"
gap 224180. .224279
/estimated_length=unknown
misc_feature 224280. .225441
/note="wgs_end_extension
clone_end:Sp6"
gap 229533. .229632
/estimated_length=unknown
gap 230666. .230765
/estimated_length=unknown
gap 231998. .232097
/estimated_length=unknown
gap 234782. .234881
/estimated_length=unknown

ORIGIN

Query Match 4.5%; Score 21; DB 12; Length 238377;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 388 TGTAAATAGTGTTTTTGAGTT 408
||||||||||||||||||||
Db 178170 TGTAAATAGTGTTTTTGAGTT 178150

Search completed: November 4, 2006, 01:20:27

Job time : 2602.63 secs

This Page Blank (uspto)